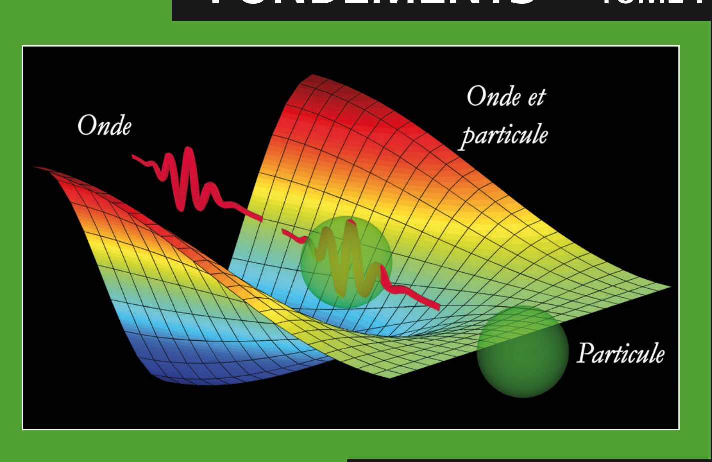
SAVOIRS PHYSIQUE ACTUELS

PHYSIQUE QUANTIQUE

FONDEMENTS – TOME I



3e ÉDITION



MICHEL LE BELLAC

PRÉFACES DE
CLAUDE COHEN-TANNOUDJI
ET DE FRANCK LALOË

CNRS ÉDITIONS
edp sciences

Michel Le Bellac

Physique quantique

Tome I : Fondements

3e édition

S A V O I R S   A C T U E L S
EDP Sciences/CNRS Éditions

Illustration de couverture : Vue d'artiste du comportement d'un photon. On observe une transition continue depuis un comportement ondulatoire (arrière-plan du dessin) à un comportement corpusculaire (avant-plan du dessin). F. Kaiser, T. Coudreau, P. Milman, D. Ostrowsky and S. Tanzilli, Entanglement enabled delayed choice experiment, Science **338***, 637 (2012). Copyright : F. Kaiser et S. Tanzilli, CNRS. Courtoisie de Sébastien Tanzilli.*

Imprimé en France.

ISBN EDP Sciences 978-2-7598-0803-8

ISBN CNRS Éditions 978-2-271-07736-3

Table des matières

Tome I : Fondements

Tome II : Applications et exercices corrigés

Avant-propos

Ce livre est issu de cours donnés à l'Université de Nice-Sophia Antipolis dans les années 1970/1980 en maîtrise de physique et en DEUG MP deuxième année, et ensuite dans les années 2000 en licence (L3) et en master de physique M1 et M2. Il est divisé en un tome 1 : "Fondements" et un tome 2 : "Applications". Les dix premiers chapitres du premier tome correspondent à un cours standard de mécanique quantique niveau L3, tandis que le chapitre 11 et le deuxième tome sont plutôt du niveau M1 et M2. Le livre contient environ 200 exercices de longueur et de difficulté variées ; plus des trois quarts de ces exercices ont été effectivement utilisés pour des séances de travaux dirigés ou des examens et une sélection de corrigés est disponible à la fin du second tome. En plus des étudiants de licence et de master et ceux des Écoles d'Ingénieurs, ce livre est susceptible d'intéresser un large public de physiciens : étudiants en thèse, chercheurs, enseignants du second degré ou du supérieur qui souhaitent trouver une introduction à la littérature récente ou simplement rafraîchir leurs connaissances en physique quantique.

L'organisation du livre diffère notablement de celle des textes classiques, qui prennent tous comme point de départ l'équation de Schrödinger, ce qui oblige à exposer les principes de base de la mécanique quantique dans un cas qui n'est pas le plus simple et a l'inconvénient de masquer ces principes par des calculs souvent fastidieux. Je me suis efforcé au contraire d'introduire les fondements de la mécanique quantique sur les exemples les plus simples et l'équation de Schrödinger apparaît seulement au chapitre 8. L'approche suivie consiste à mener jusqu'à à son terme la logique qu'avait adoptée Feynman (Feynman *et al.* [1965]) : développer au maximum un traitement algébrique et exploiter les symétries, en présentant la mécanique quantique dans son cadre autonome, sans faire référence à la physique classique. Cette logique a de nombreux avantages.

- L'approche algébrique permet de traiter des problèmes simples dans des espaces de dimension finie, par exemple de dimension deux : polarisation d'un photon, spin 1/2, atome à deux niveaux...

- Elle permet d'énoncer de la façon la plus claire les postulats de la mécanique quantique, en séparant ce qui est fondamental de ce qui ne l'est pas (par exemple le principe de correspondance n'est pas un postulat fondamental).
- L'exploitation des propriétés de symétrie permet l'introduction la plus générale des propriétés physiques fondamentales : impulsion, moment angulaire… comme générateurs infinitésimaux de ces symétries, sans faire appel au principe de correspondance et à un analogue classique.

Les aspects pédagogiques ont fait l'objet d'une attention particulière. La progression des chapitres a été soigneusement étudiée, les chapitres 2 à 5 utilisant uniquement des espaces de dimension finie ; c'est seulement une fois les bases acquises que l'on passe au cas général à partir du chapitre 6. Un effort a été fait sur le vocabulaire, afin d'éviter certaines expressions historiquement datées et qui peuvent être un obstacle à la compréhension de la mécanique quantique : suivant la modernisation du vocabulaire préconisée par Lévy-Leblond [1999], "propriété physique" est utilisé au lieu d'"observable", "inégalité de Heisenberg" au lieu de "principe d'incertitude", des expressions comme "complémentarité" ou "dualité onde-corpuscule" ont été evitées, etc.

Les chapitres clés du livre, c'est-à-dire ceux qui divergent de la façon la plus évidente de l'exposé traditionnel, sont les chapitres 3, 4, 5 et 7. Le chapitre 3 introduit l'espace des états sur l'exemple de la polarisation des photons et montre comment passer d'une amplitude ondulatoire à une amplitude de probabilité. Le spin 1/2 permet ensuite d'initier le lecteur à un problème sans analogue classique et à ses propriétés essentielles : algèbre des matrices de Pauli, matrices de rotation, précession de Larmor. Dans le chapitre 4, la distinction entre le cadre conceptuel général de la mécanique quantique et la modélisation d'un problème concret est soigneusement expliquée. Le chapitre 5 met en pratique la mécanique quantique sur des applications simples et physiquement importantes, dans le cas de systèmes dont le nombre de niveaux est fini. Il introduit l'interaction d'un système à deux niveaux atomique ou moléculaire avec un champ électromagnétique oscillant et développe des applications importantes : RMN, émission et absorption de photons, lasers, horloge atomiques… Le chapitre 7 a comme objectif l'étude des symétries à partir du théorème de Wigner, qui est généralement ignoré des manuels malgré son importance cruciale. La symétrie de rotation permet de définir le moment angulaire comme générateur infinitésimal et de démontrer immédiatement les relations de commutation de $\vec{J}$ en soulignant leur origine géométrique.

Le chapitre 11 sur l'intrication et la non-localité quantiques et les chapitres du deuxième tome pourront servir d'introduction à des sujets importants, et dont beaucoup sont apparus récemment : états intriqués, décohérence, information quantique, états du champ électromagnétique, intégrale de chemin, manipulation d'atomes par laser, condensats de Bose-Einstein, équations pilotes pour les systèmes ouverts et enfin mécanique quantique relativiste (équation de Dirac). Cette nouvelle édition, contrairement à la première qui

était avant tout un livre d'enseignement, couvre une trop grande variété de domaines pour qu'elle puisse encore correspondre à un cours d'une année. Le guide de lecture ci-dessous devrait permettre au lecteur de s'orienter dans le livre pour un cours d'introduction à la mécanique quantique.

Guide de lecture

Seules les sections 1.3 à 1.5 du chapitre 1 sont indispensables pour la suite, et le chapitre 2 peut être omis par le lecteur qui possède le niveau L2 en algèbre. Les chapitres 3 et 5 constituent, à mon avis, le cœur d'une introduction à la mécanique quantique : les principes de base sont introduits sur les exemples de la polarisation du photon et du spin 1/2, et le rôle de la symétrie de rotation est mis en valeur dans le cas du spin 1/2. Des applications simples de ces principes de base sont exposées dans le chapitre 5 : résolution de l'équation de Schrödinger indépendante du temps pour des problèmes de chimie quantique élémentaire, oscillations de Rabi, le tout illustré par la RMN, l'émission et l'absorption de photons par un atome à deux niveaux, le laser et les horloges atomiques. Le chapitre 4 (Postulats de la mécanique quantique) est un peu plus abstrait : le lecteur pourra se contenter d'une première prise de contact, et y revenir après avoir acquis un peu plus de familiarité avec la mécanique quantique. Dans le chapitre 6, le lecteur peut se limiter à la section 6.3, qui donne les recettes "mathématiques" pour la suite. Les chapitres 7 à 10 sont classiques dans tout cours de mécanique quantique niveau L3. Toutefois la section 7.4 (invariance galiléenne) peut être omise sans dommage pour la suite et la section 8.4 (bandes d'énergie) peut trouver sa place dans un cours de physique du solide. Il en est de même pour la section 10.3 (niveaux de Landau) qui peut aussi faire partie d'un cours de physique du solide. Dans une première lecture du chapitre 11, on peut se limiter à l'introduction de l'opérateur statistique qui sera indispensable pour la suite. Il serait toutefois dommage de ne pas consacrer un peu de temps à la présentation élémentaire des inégalités de Bell (section 11.2). Le lecteur peut aborder les autres sections de ce chapitre, en particulier celle sur l'information quantique, en fonction de ses intérêts. Il en est de même pour les chapitres du tome 2 qui sont indépendants en première approximation ; le lecteur familier du tome 1 peut les aborder dans un ordre indifférent. Le chapitre 12 (Méthodes semi-classiques) est entièrement optionnel, et ses quatre grandes parties : intégrale de chemin, approximation BKW, distribution de Wigner et phases géométriques sont largement indépendantes. Dans une première lecture, il est possible de se limiter aux sections 14.1 et 14.2 du chapitre 14 (Particules identiques). Enfin les sections 15.1 et 15.2 sont une introduction standard à la physique atomique, les autres sections de ce chapitre étant optionnelles.

Remerciements. J'ai bénéficié des critiques et suggestions de Pascal Baldi, Jean-Pierre Farges, Yves Gabellini, Thierry Grandou, Jacques Joffrin, Jean-Marc Lévy-Leblond, Christian Miniatura et tout particulièrement de Michel Brune, Jean Dalibard, Franck Laloë, Fabrice Mortessagne, Jean-Pierre

Romagnan, François Rocca et Sébastien Tanzilli qui ont lu de larges extraits du manuscrit. Je remercie également David Wilkowski qui a inspiré le texte de plusieurs exercices du chapitre 15. Je suis bien entendu entièrement responsable du texte final. L'aide de Karim Bernardet et de Fabrice Mortessagne a été décisive dans la réalisation des figures, et je tiens à remercier Christian Taggiasco pour sa compétence et sa disponibilité dans l'installation et la maintenance de l'ensemble des logiciels nécessaires. Enfin ce livre n'aurait pas vu le jour sans les encouragements et le soutien sans faille de Michèle Leduc, et je suis très reconnaissant à Claude Cohen-Tannoudji et Franck Laloë qui ont bien voulu le préfacer.

N.B. Cet ouvrage utilise le point décimal.

Préface de la première édition

La naissance de la physique quantique date d'un siècle et cette description des phénomènes physiques, qui a transformé notre vision du monde, n'est toujours pas remise en cause, ce qui est exceptionnel pour une théorie scientifique. Ses prédictions ont toujours été vérifiées par l'expérience avec une précision impressionnante. Les concepts fondamentaux, comme les amplitudes de probabilité, les superpositions linéaires d'états, qui semblent si étranges pour notre intuition quand on les rencontre pour la première fois, restent toujours essentiels. Une évolution importante s'est toutefois manifestée au cours des dernières décennies. Les progrès spectaculaires des techniques d'observation, des méthodes de manipulation des atomes, permettent maintenant de réaliser des expériences si délicates qu'elle n'étaient considérées que comme des "expériences de pensée" par les pères fondateurs de la mécanique quantique. L'existence des corrélations quantiques "non séparables", qui est à la base du "paradoxe" de Einstein-Podolsky-Rosen et qui violent les fameuses inégalités de Bell, a pu être confirmée expérimentalement avec une grande précision. Les états "intriqués" de deux systèmes, qui manifestent de telles corrélations quantiques, sont mieux compris, et même utilisés pour des applications concrètes, comme la cryptographie quantique. L'intrication d'un appareil de mesure avec son environnement se révèle une piste intéressante pour une meilleure compréhension du processus de mesure.

Parallèlement à ces progrès conceptuels, on assiste également à une invasion de notre monde quotidien par des dispositifs dont le principe de fonctionnement repose sur des phénomènes quantiques. Les sources laser qui sont utilisées pour la lecture des disques compacts, l'ophtalmologie ou les télécommunications optiques, sont basées sur l'amplification de la lumière par des systèmes atomiques dont les populations sont inversées. La résonance magnétique des noyaux des atomes est couramment utilisée dans les hôpitaux pour prendre des images de plus en plus précises des organes du corps humain. Des millions de transistors sont inclus dans les puces qui permettent à nos ordinateurs d'effectuer des opérations à des vitesses prodigieuses.

Il est donc clair qu'un enseignement moderne de la physique quantique doit tenir compte de ces développements, pour donner à l'étudiant ou au chercheur qui désire s'instruire une image plus précise des progrès réalisés et pour accroître sa motivation de mieux comprendre les phénomènes physiques dont l'importance conceptuelle et pratique est de plus en plus évidente. C'est ce défi qu'essaie de relever avec succès Michel Le Bellac dans le présent ouvrage.

Chacun des 14 chapitres de ce livre contient en effet, en plus d'un exposé clair et concis des notions de base, de nombreuses discussions présentant des développements conceptuels ou expérimentaux récents, qui permettent au lecteur de se faire une idée précise des avancées de la discipline et de ses grandes tendances d'évolution. Le chapitre 6 sur les états intriqués est bien caractéristique d'un tel choix de présentation. Au lieu de mettre l'accent sur les propriétés mathématiques du produit tensoriel de deux espaces d'états, ce qui est un peu austère et rébarbatif, ce chapitre préfère centrer la discussion sur la notion d'intrication, et introduire de nombreux exemples de développements théoriques et expérimentaux (dont certains sont très récents) : inégalités de Bell, tests de ces inégalités, en particulier les plus récents utilisant la conversion paramétrique, les états GHZ (Greenberger, Horne, Zeilinger), la notion de décohérence illustrée par des expériences modernes d'électrodynamique en cavité, et qui sera reprise plus en détail dans une annexe, la téléportation. Comme on le voit, il est difficile d'imaginer une immersion plus complète dans l'un des domaines les plus actifs actuellement de la physique quantique. De nombreux exemples de présentation moderne peuvent être donnés à propos d'autres chapitres : interférences d'ondes de de Broglie réalisées avec des neutrons lents, ou des atomes refroidis par laser ; microscopie à effet tunnel ; fluctuations du champ quantique et effet Casimir ; transformations de jauge non abéliennes ; équations de Bloch optiques ; forces radiatives exercées par des faisceaux laser sur des atomes ; piège magnéto-optique ; oscillations de Rabi dans le vide d'une cavité, etc.

Je suis vraiment admiratif devant l'effort fait par l'auteur pour donner à son lecteur une vision si moderne et si attrayante de la physique quantique. Certes, les développements décrits ne peuvent pas toujours être analysés en grand détail, et le lecteur devra fournir un effort personnel pour parvenir à une compréhension plus approfondie du sujet étudié. Il sera aidé en cela par la bibliographie détaillée qu'il trouvera, soit au cours du chapitre sous forme de note de bas de page, soit à la fin de chaque chapitre. Je suis convaincu qu'un tel ouvrage permettra une meilleure compréhension de la physique quantique et stimulera un plus grand intérêt pour cette discipline aussi centrale. Je remercie Michel Le Bellac pour cette contribution importante qui va certainement donner une image plus vivante de la physique.

Claude Cohen-Tannoudji

Préface de la troisième édition

Les éditions précédentes du livre de Michel Le Bellac ont rencontré un succès mérité, car elles concilient de façon exceptionnelle clarté et rigueur sans pour autant sacrifier la concision. Donner autant d'information utile dans un nombre de pages raisonnable rend le livre particulièrement agréable à utiliser. La toute dernière édition augmente encore la portée de l'ouvrage, puisqu'elle bénéficie d'additions et de réorganisations qui modifient sensiblement le développement des chapitres. L'intrication et la non-localité quantiques font maintenant l'objet d'un chapitre particulier très riche couvrant, non seulement divers aspects du théorème de Bell, mais également les relations entre décohérence et mesure et l'information quantique. Des exercices corrigés originaux complètent encore l'information, comme par exemple celui qui discute la réalisation de portes quantiques avec des ions. De même, la quantification du champ électromagnétique est maintenant promue au niveau d'un chapitre entier (alors qu'antérieurement, elle était une partie du chapitre sur l'oscillateur harmonique) ; divers aspects sont traités, comme l'effet d'une lame semi-réfléchissante sur les modes du champ, ou encore l'hamiltonien de Jaynes-Cummings. La théorie de la diffusion est maintenant enrichie d'une partie formelle et opératorielle (équation de Lippmann-Schwinger, matrice T, etc.).

Une addition notable est celle d'un chapitre 19, "Physique quantique relativiste". Une première partie est consacrée à l'analyse de Wigner des propriétés universelles d'un système quantique élémentaire déduites des symétries du groupe de Poincaré. C'est important car cela permet de comprendre l'origine profonde de la masse et du spin en termes de symétries fondamentales. Il est curieux que relativement peu d'ouvrages d'enseignement traitent ce sujet. Cette analyse conduit à l'équation de Dirac, dont les propriétés sont discutées. La quantification de l'équation de Klein-Gordon et de l'équation de Dirac, considérées comme des équations classiques, donne un aperçu utile sur la théorie relativiste des champs.

À tout ceci s'ajoute un grand nombre d'exercices très riches munis de leurs corrigés, parmi lesquels on peut citer les forces de Van des Waals et l'effet Casimir, la précession de Thomas, le modèle en couches des noyaux, l'étude des expériences de création d'états de Fock pour des photons en cavité, la capture radiative de neutrons par le proton, la hiérarchie BBGKY, les équations pilotes (Lindblad, Fokker-Planck), etc. De nombreuses figures illustrent le texte d'une façon qui en rend l'accès plus agréable au lecteur.

Bref, plus encore dans cette nouvelle version, cet ouvrage est un concentré d'informations et d'idées dans presque tous les domaines qui touchent au quantique, qui servira beaucoup non seulement aux étudiants, mais également aux physiciens des laboratoires comme ouvrage de référence.

Franck Laloë

Chapitre 1

Introduction

Le premier objectif de ce chapitre est d'exposer succinctement quelques notions de base sur l'organisation de la matière, en reprenant et en précisant les acquis de cours de physique (et de chimie) antérieurs, et en particulier les notions de physique microscopique ; il s'agira d'un survol, et la grande majorité des énoncés seront donnés sans démonstration et sans discussion détaillée. Le deuxième objectif est de décrire brièvement quelques étapes cruciales des débuts de la physique quantique ; nous ne suivrons ni l'ordre historique strict, ni les arguments qu'utilisèrent au début du siècle dernier les pères fondateurs de la mécanique quantique, mais nous insisterons plutôt sur les concepts qui nous serviront par la suite. Le troisième objectif est d'introduire des notions de base, comme celles de particule quantique ou de niveau d'énergie, qui reviendront de façon récurrente tout au long du livre. Nous nous appuierons sur la théorie de Bohr, qui permet d'expliquer de façon simple, sinon convaincante, la notion de quantification des niveaux d'énergie et le spectre de l'atome d'hydrogène. Ce chapitre est à relire ultérieurement, lorsque les bases de la mécanique quantique auront été explicitées et illustrées par des exemples. D'un point de vue pratique, il est possible de sauter en première lecture les considérations générales des sections 1 et 2 et de commencer ce chapitre par la section 3, quitte à revenir ultérieurement aux deux premières sections lorsqu'il sera fait appel aux notions qui y ont été introduites.

1.1 Structure de la matière

1.1.1 Échelles de longueur : de la cosmologie aux particules élémentaires

Le tableau 1.1 donne l'ordre de grandeur en mètres des dimensions de quelques objets typiques, en partant des dimensions de l'Univers pour descendre à celles de la physique subatomique. Une unité de longueur

Tab. 1.1 – Ordres de grandeur de quelques distances typiques en mètres.

Univers connu	Rayon de la Galaxie	Distance Terre-Soleil	Rayon de la Terre	Homme	Insecte
1.3×10^{26}	$\simeq 5 \times 10^{20}$	1.5×10^{11}	6.4×10^{6}	$\simeq 1.7$	0.01 à 0.001
bactérie *E. coli*	virus HIV	fullerène C_{60}	atome	noyau de plomb	proton
$\simeq 2 \times 10^{-6}$	1.1×10^{-7}	0.7×10^{-9}	$\simeq 10^{-10}$	7×10^{-15}	0.8×10^{-15}

commode pour les distances astrophysiques est l'année-lumière : 1 année-lumière $= 0.95 \times 10^{16}$ m. Les sous-multiples du mètre utiles en physique microscopique sont le micromètre : 1 μm $= 10^{-6}$ m, le nanomètre : 1 nm $= 10^{-9}$ m, et le femtomètre (ou fermi) : 1 fm $= 10^{-15}$ m. L'exploration des objets à l'échelle microscopique se fait souvent à l'aide d'un rayonnement électromagnétique[1] dont la longueur d'onde est de l'ordre de grandeur des dimensions caractéristiques de l'objet à étudier (observation au microscope, aux rayons X...). On sait en effet que la limite de résolution est fixée par la longueur d'onde utilisée : quelques fractions de μm pour un microscope utilisant de la lumière visible, quelques fractions de nanomètre pour des rayons X. La gamme des longueurs d'onde du rayonnement électromagnétique (infra-rouge, visible, ...) est résumée sur la figure 1.1.

1.1.2 États de la matière

Nous serons particulièrement intéressés par les phénomènes à l'échelle microscopique, et il est utile de rappeler quelques notions élémentaires sur la description microscopique de la matière. La matière peut se présenter sous deux formes : une forme ordonnée, le solide cristallin, et une forme non ordonnée : liquide, gaz, solide amorphe.

Le solide cristallin présente un *ordre à longue distance*. La figure 1.2 donne l'exemple de la structure microscopique du chlorure de sodium : on constate que le motif du cristal se répète avec une périodicité $l = 0.56$ nm, le *pas du réseau*. Partant d'un ion chlore ou d'un ion sodium, et suivant une des arêtes de la structure cubique, on retrouvera un ion chlore ou un ion sodium à une distance $n \times 0.56$ nm où n est un nombre entier : c'est ce que l'on appelle un ordre à longue distance.

Les liquides, les gaz et les solides amorphes ne possèdent pas d'ordre à longue distance. Prenons l'exemple d'un liquide monoatomique, celui de l'argon liquide. En première approximation, les atomes d'argon peuvent être représentés par des sphères impénétrables de diamètre $\sigma \simeq 0.36$ nm. La

1. D'autres techniques d'exploration sont la diffusion de neutrons (exercice 1.6.4), la microscopie électronique, la microscopie à effet tunnel (§ 12.4.4), etc.

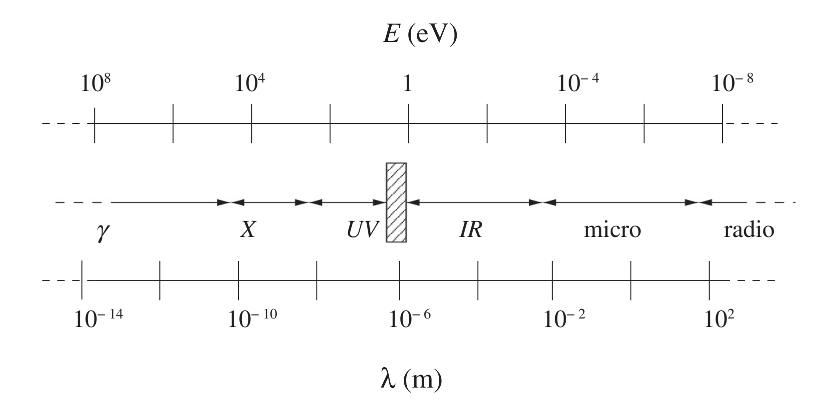


Fig. 1.1 – Longueur d'onde de rayonnements électromagnétiques et énergie des photons correspondants. La boîte hachurée représente le domaine visible. Les frontières entre les différents types de rayonnement (par exemple la frontière entre les rayons γ et les rayons X) ne sont pas définies de façon stricte. Un photon d'énergie $E = 1$ eV a une longueur d'onde $\lambda = 1.24 \times 10^{-6}$ m, une fréquence $\nu = 2.42 \times 10^{14}$ Hz et une fréquence angulaire $\omega = 1.52 \times 10^{15}$ rad.s^{-1}.

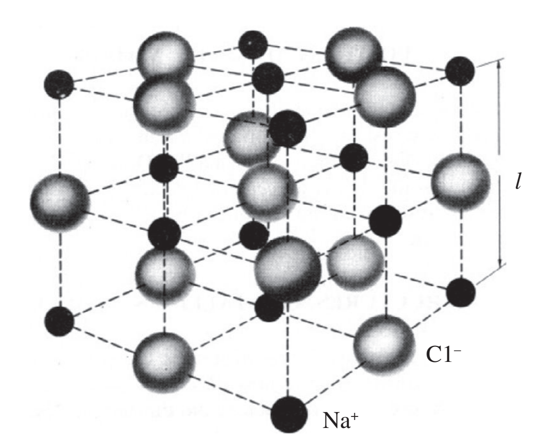


Fig. 1.2 – Arrangement des atomes dans le cristal de chlorure de sodium. Les ions chlore Cl^- sont plus gros que les ions sodium Na^+.

figure 1.3 représente schématiquement une configuration des atomes pour un liquide où les sphères sont pratiquement en contact, mais sont disposées de façon désordonnée. Partant du centre d'un atome pris comme origine, la probabilité $\mathsf{p}(r)$ de trouver le centre d'un autre atome à une distance r de ce centre sera pratiquement nulle tant que $r \lesssim \sigma$. Cette probabilité deviendra au contraire maximale pour $r = \sigma, 2\sigma, \ldots$ et oscillera ensuite pour se stabiliser à une constante, alors que dans le cas d'un solide cristallin, la fonction $\mathsf{p}(r)$

présente des pics, quelle que soit la distance à l'origine. L'argon gazeux possède le même type de configuration que le liquide, la seule différence étant que les atomes sont beaucoup plus éloignés les uns des autres. Toutefois, la différence entre gaz et liquide disparaît au point critique, et on peut passer continûment du gaz au liquide ou inversement en contournant le point critique, alors qu'un tel passage continu est impossible vers un solide, parce que le type d'ordre est qualitativement différent.

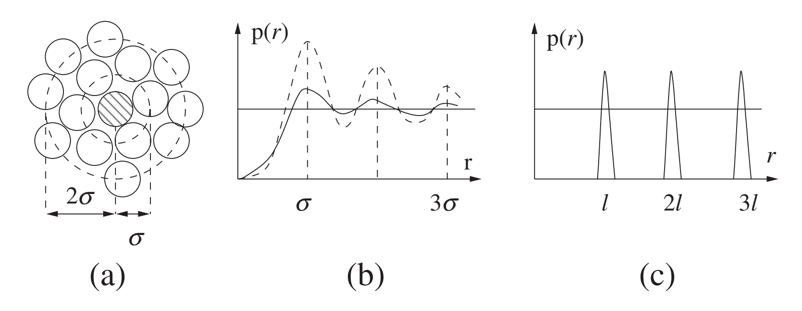


Fig. 1.3 – (a) Arrangement des atomes dans l'argon liquide. (b) Probabilité $\mathsf{p}(r)$ pour un liquide (tirets) et pour un gaz (trait continu). (c) Probabilité $\mathsf{p}(r)$ pour un cristal simple.

Nous avons pris comme exemple un gaz monoatomique, mais en général la brique de base est une combinaison d'atomes, la molécule : N_2, O_2, H_2O, etc. Certaines molécules comme les protéines peuvent contenir des milliers d'atomes ; par exemple le poids moléculaire de l'hémoglobine est de l'ordre de 64 000. Une réaction chimique est un réarrangement d'atomes : les atomes des molécules initiales se redistribuent pour donner les molécules finales

$$H_2 + Cl_2 \rightarrow 2HCl$$

L'atome est composé d'un noyau atomique (ou simplement noyau) chargé positivement et d'électrons chargés négativement. Plus de 99.9 % de la masse de l'atome se trouve dans le noyau atomique, car le rapport de la masse de l'électron m_e à celle du proton m_p est $m_e/m_p \simeq 1/1836$. L'atome est de dix à cent mille fois plus gros que le noyau : la dimension typique d'un atome[2] est 1 Å (1 Å $= 10^{-10}$ m $= 0.1$ nm), celle d'un noyau de quelques fermis (ou femtomètres).

Un noyau atomique est formé de protons et de neutrons, les premiers chargés électriquement et les seconds neutres ; les masses du proton et du neutron sont identiques à 0.1 % près, et on pourra souvent négliger cette différence de masse. Le *numéro atomique* Z est le nombre de protons du noyau, et aussi le nombre d'électrons de l'atome correspondant, de façon à assurer

2. Il nous arrivera d'utiliser l'Angström, qui est typique de la dimension atomique, plutôt que le nm. Par un hasard (?) heureux, le symbole fm peut aussi bien désigner le femtomètre que le fermi, unité de longueur des physiciens nucléaires.

sa neutralité électrique. Le *nombre de masse* A est le nombre de protons plus le nombre de neutrons N : $A = Z + N$. Les protons et les neutrons sont appelés collectivement *nucléons*. Les réactions nucléaires sont pour les protons et les neutrons l'analogue des réactions chimiques pour les atomes : une réaction nucléaire est une redistribution des protons et des neutrons dans des noyaux différents des noyaux initiaux, de même qu'une réaction chimique est une redistribution des atomes dans des molécules différentes des molécules initiales. Un exemple de réaction nucléaire est la réaction de fusion des noyaux de deutérium (^{2}H : un proton + un neutron) et de tritium (^{3}H : un proton + deux neutrons) pour donner un noyau d'hélium[4] (^{4}He : deux protons + deux neutrons) et un neutron libre

$$^2\mathrm{H} + {}^3\mathrm{H} \to {}^4\mathrm{He} + n + 17.6\ \mathrm{MeV}$$

La réaction dégage une énergie de 17.6 MeV et pourrait être utilisée dans un futur (probablement lointain) pour produire de l'énergie à grande échelle, l'énergie de fusion. Le réacteur de recherche ITER en construction à Cadarache est peut-être une première étape dans cette direction.

Dans la composition de l'atome en noyau et électrons, de même que dans celle du noyau en protons et neutrons, un concept important est celui *d'énergie de liaison*. Considérons un objet stable C formé de deux objets A et B : C est appelé *état lié* de A et B. La désintégration $C \to A + B$ ne sera pas possible si la masse m_C de C est inférieure à la somme des masses m_A et m_B de A et de B, c'est-à-dire si l'énergie de liaison[3] E_L

$$E_L = (m_A + m_B - m_C)c^2 \tag{1.1}$$

est positive ; c est la vitesse de la lumière. E_L est l'énergie qu'il faut fournir pour dissocier C en $A + B$. En physique atomique, cette énergie est appelée *énergie d'ionisation* : c'est l'énergie nécessaire pour dissocier un atome en un ion positif et un électron, ou, en d'autres termes, pour arracher un électron à l'atome. Pour les molécules, E_L est l'énergie de dissociation, l'énergie nécessaire pour dissocier la molécule en atomes. Une particule ou un noyau instable dans une certaine configuration peut parfaitement être stable dans une autre configuration. Le neutron libre (n), par exemple, est instable : en un temps d'une quinzaine minutes en moyenne (plus précisément 880 s), il se désintègre en un proton (p), un électron (e) et un antineutrino électronique ($\overline{\nu}_e$), ce qui correspond à la désintégration de base de la radioactivité β

$$n^0 \to p^+ + e^- + \overline{\nu}_e^0 \tag{1.2}$$

3. En raison de la célèbre relation d'Einstein $E = mc^2$, ou tout simplement par analyse dimensionnelle, on peut relier masse et énergie et exprimer par exemple les masses en J$/c^2$ ou en eV$/c^2$.

où nous avons indiqué en exposant la charge des particules en unités de la charge du proton. Cette désintégration est possible, car les masses[4] des particules dans (1.2) vérifient

$$m_n c^2 > (m_p + m_e + m_{\overline{\nu}})c^2$$

$$m_n \simeq 939.5\,\mathrm{MeV/c}^2 \quad m_p \simeq 938.3\,\mathrm{MeV/c}^2 \quad m_e \simeq 0.51\,\mathrm{MeV/c}^2 \quad m_{\overline{\nu}_e} \simeq 0$$

En revanche le neutron ne se désintègre pas dans les noyaux atomiques stables, par exemple dans le noyau deutérium, ou deutéron, (^{2}H), car

$$m_{^2\mathrm{H}} \simeq 1875.6\ \mathrm{Mev/c}^2 < (2m_p + m_e + m_{\overline{\nu}_e}) \simeq 1878.3\,\mathrm{MeV/c}^2$$

et la désintégration

$$^2\mathrm{H} \to 2p + e + \overline{\nu}_e$$

est impossible : le deutéron est un état lié proton-neutron.

1.1.3 Constituants élémentaires

Nous avons décomposé les molécules en atomes, les atomes en électrons et noyaux, les noyaux en protons et neutrons. Peut-on aller encore plus loin, par exemple décomposer le proton ou l'électron en constituants plus élémentaires ? Peut-on dire par exemple à partir de (1.2) que le neutron est "formé" d'un proton, d'un électron et d'un antineutrino ? Un argument simple fondé sur les inégalités de Heisenberg montre que l'électron ne peut pas préexister dans le neutron (exercice 8.6.4), et qu'il est *créé* au moment de la désintégration : on ne peut donc pas dire que le neutron est "composé" d'un proton, d'un électron et d'un neutrino. On pourrait aussi penser à "casser" un proton ou un neutron en des composants plus élémentaires, en les bombardant par des particules énergiques, et répéter ce qui se passe par exemple quand on bombarde un deutéron avec des électrons de quelques MeV

$$e + {}^2\mathrm{H} \to e + p + n$$

Le deutéron ^{2}H a été cassé en ses constituants, un proton et un neutron. L'histoire ne se répète pas lorsque l'on bombarde un proton avec des électrons. Pour des électrons peu énergiques, les collisions sont élastiques

$$e + p \to e + p$$

mais pour des électrons d'énergie suffisante (quelques centaines de MeV), au lieu de casser le proton, on *crée* d'autres particules, par exemple dans des

4. Trois expériences : Fukuda *et al.* [2001], Ahmad *et al.* [2001] et Eguchi *et al.* [2003] montrent de façon convaincante que la masse du neutrino est non nulle, probablement de l'ordre de 10^{-2} eV/c^2 : *cf.* l'exercice 4.4.6 sur les oscillations neutrino. Pour une revue, voir par exemple Wark [2005].

réactions du type

$$e + p \to e + p + \pi^0$$
$$e + p \to e + n + \pi^+ + \pi^0$$
$$e + p \to e + K^+ + \Lambda^0$$

où les mésons π et K et l'hypéron Λ^0 sont de nouvelles particules, dont la nature importe peu ici. Le point crucial est qu'elles ne sont pas présentes initialement dans le proton, mais qu'elles sont créées au moment de la réaction.

Il arrive donc un moment où il ne semble plus possible de décomposer la matière en constituants de plus en plus élémentaires. On peut alors se poser la question suivante : quel est le critère d'élémentarité ? Le point de vue actuel est d'appeler élémentaires les particules qui apparaissent comme ponctuelles dans leurs interactions avec d'autres particules. Avec ce point de vue, l'électron, le neutrino et le photon sont élémentaires, tandis que le proton et le neutron sont "composés" de quarks : les guillemets sont importants, car les quarks n'existent pas à l'état libre[5], et cette "composition" du proton en quarks est tout à fait différente de la composition du deutéron en proton et neutron. Il existe seulement des preuves indirectes (mais convaincantes !) de cette composition en quarks.

Dans l'état actuel de nos connaissances[6], il existe trois familles de particules élémentaires, ou "particules de matière" de spin 1/2[7]. Le tableau 1.2 en donne la liste ; la charge électrique q est exprimée en unités de la charge du proton. Chaque famille se compose de leptons et de quarks, et à chaque particule correspond une antiparticule de charge opposée. Les leptons de la première famille sont l'électron et son antiparticule le positron e^+, ainsi que le neutrino électronique ν_e et son antiparticule l'antineutrino électronique $\overline{\nu}_e$, et les quarks de cette famille sont le quark "up" (u) de charge 2/3 et le quark "down" (d) de charge −1/3, avec bien sûr les antiquarks $\overline{u}$ et $\overline{d}$ correspondants, de charge −2/3 et 1/3, respectivement. Le proton est une combinaison uud et le neutron une combinaison udd. Cette première famille suffit à notre vie courante, puisqu'elle permet de fabriquer la matière ordinaire, le neutrino étant indispensable dans le cycle des réactions nucléaires assurant la bonne marche du Soleil. Si l'existence de la première famille peut se justifier par un argument anthropocentrique : si elle n'existait pas, nous ne serions pas là pour en parler ! la raison d'être des deux autres familles reste aujourd'hui tout à fait mystérieuse[8]. Enfin, le "boson de Higgs" est la particule

5. Ce qui fait que la notion de "masse" d'un quark est une notion complexe, du moins pour les quarks dits "légers", up, down et étrange. On retrouve une masse (presque) usuelle pour les quarks lourds b et t.

6. Il existe un argument très fort pour limiter à trois ce nombre de familles : des expériences au CERN ont montré en 1992 que le nombre de familles était limité à trois, à condition que les neutrinos aient une masse inférieure à 45 GeV/c^2. La valeur expérimentale actuelle du nombre de familles est de 2.984 ± 0.008.

7. Le spin 1/2 est défini au chapitre 3 et le spin en général au chapitre 10.

8. *Cf.* la célèbre interrogation de Rabi : "Who ordered the muon ?". Cependant, on sait que chaque famille doit être complète : ainsi l'existence du quark top et sa masse ont été

TAB. 1.2 – Les particules de matière. Les charges électriques sont mesurées en unités de la charge du proton.

	lepton $q = -1$	neutrino $q = 0$	quark $q = 2/3$	quark $q = -1/3$
famille 1	électron	neutrino$_e$	quark up	quark down
famille 2	muon	neutrino$_\mu$	quark charmé	quark étrange
famille 3	tau	neutrino$_\tau$	quark top	quark b

responsable des masses de toutes les autres particules massives, en dehors de celles des neutrinos. Il est très vraisemblable que ce boson de Higgs a été identifié en juillet 2012 au collisionneur LHC du CERN à Genève, avec une masse de 125 GeV/c^2.

À ces particules, il faut ajouter celles qui “transportent les interactions” : le photon pour les interactions électromagnétiques, les bosons W et Z pour les interactions faibles, les gluons pour les interactions fortes et le graviton pour les interactions gravitationnelles[9], ce qui nous amène naturellement à introduire ces interactions.

1.1.4 Interactions (ou forces) fondamentales

On distingue quatre[10] types d'interactions (ou de forces) fondamentales : fortes, électromagnétiques, faibles et gravitationnelles. Les *interactions électromagnétiques* sont celles qui vont jouer un rôle majeur dans ce livre : ce sont elles qui régissent le comportement des atomes, des molécules, des solides, etc. Les forces électriques de la loi de Coulomb sont dominantes : rappelons que si une charge q est immobile à l'origine des coordonnées, la force qu'elle exerce sur une charge immobile q' située en un point $\vec{r}$ est

$$\vec{F} = \frac{qq'}{4\pi\varepsilon_0}\,\frac{\hat{r}}{r^2} \tag{1.3}$$

prédites plusieurs années avant sa découverte expérimentale en 1994. En raison de sa masse élevée, environ 180 fois la masse du proton, il a fallu attendre la construction du Tevatron aux USA pour le produire.

9. En toute rigueur, les interactions électromagnétiques et les interactions faibles sont maintenant unifiées en interactions électrofaibles ; les gluons, tout comme les quarks, n'existent pas à l'état libre. Enfin, l'existence du graviton est pour le moment hypothétique : on n'a même pas encore détecté sur la Terre sa version classique, les ondes gravitationnelles.

10. La “cinquième force” revient périodiquement sur le devant de la scène, mais elle disparaît tout aussi périodiquement !

où $\hat{r}$ est le vecteur unitaire[11] $\vec{r}/r$, $r = |\vec{r}|$ et ε_0 la permittivité du vide. Si les charges sont en mouvement avec une vitesse v, on doit aussi tenir compte des forces magnétiques, mais celles-ci sont plus faibles que la force de Coulomb par un facteur $\sim (v/c)^2$. Pour les électrons des couches externes, $(v/c)^2 \sim (1/137)^2 \ll 1$, mais, compte tenu de la très grande précision des expériences de physique atomique, les forces magnétiques sont facilement mises en évidence dans des phénomènes comme la structure fine ou l'effet Zeeman (§ 15.2.3). La force de Coulomb (1.3) est caractérisée par

- la forme de la loi de force en $1/r^2$: la loi de force est dite *à longue portée* ;
- l'intensité de la force mesurée par une *constante de couplage* $qq'/(4\pi\varepsilon_0)$.

Le point de vue moderne, celui de la théorie quantique des champs, est que les forces électromagnétiques sont dues à l'échange de photons "virtuels"[12] entre particules chargées. La théorie quantique des champs résulte du mariage (conflictuel[13] !) entre la mécanique quantique et la relativité restreinte. Les interactions entre atomes ou molécules sont représentées par des forces effectives, par exemple les forces de van der Waals (exercice 15.6.1). Ces forces n'ont pas de caractère fondamental car elles se déduisent de la force de Coulomb : c'est le déguisement sous lequel apparaît la force de Coulomb pour des sytèmes complexes électriquement neutres.

Les *interactions fortes* sont responsables de la cohésion du noyau atomique. Contrairement à la force de Coulomb, elles décroissent exponentiellement en fonction de la distance, suivant une loi en $\exp(-r/r_0)/r^2$, avec $r_0 \simeq 1\,\mathrm{fm}$: on dit que ce sont des forces *à courte portée*. Pour $r \lesssim r_0$, elles sont très intenses. Il en résulte que les énergies caractéristiques dans un noyau sont de l'ordre du MeV, alors qu'elles sont de l'ordre de l'eV dans un atome pour les électrons des couches externes. En réalité, les forces entre nucléons ne sont pas des forces de type fondamental, car les nucléons sont, on l'a vu, des particules composées. Les forces entre nucléons sont l'analogue des forces de van der Waals pour les atomes, et les forces fondamentales sont les forces entre quarks. Cependant, la relation quantitative qui relie les forces entre nucléons et les forces entre quarks est encore loin d'être comprise. Le gluon, particule de masse nulle et de spin 1 comme le photon, joue, pour les interactions fortes, le rôle que joue le photon pour les interactions électromagnétiques. La charge est remplacée par une propriété appelée par convention couleur,

11. Nous utiliserons systématiquement la notation $\hat{r}$, $\hat{n}$, $\hat{p}$, etc. pour les vecteurs unitaires de l'espace ordinaire.

12. Ce terme sera expliqué au § 4.2.4.

13. La combinaison de la mécanique quantique et de la relativité restreinte conduit à des résultats infinis, et ces infinités doivent être contrôlées par une procédure appelée renormalisation, qui n'a vraiment été comprise et justifiée que dans les années 1970.

et la théorie des interactions fortes est appelée pour cette raison la *chromodynamique*.

Les *interactions faibles* sont responsables de la radioactivité β

$$(Z, N) \to (Z + 1, N - 1) + e^- + \overline{\nu}_e \tag{1.4}$$

dont un cas particulier est (1.2) qui s'écrit avec les notations de (1.4)

$$(0, 1) \to (1, 0) + e^- + \overline{\nu}_e$$

De même que les interactions fortes, les interactions faibles sont à courte portée, mais, comme leur nom l'indique, elles sont beaucoup moins intenses. Les vecteurs des interactions faibles sont les bosons de spin 1 chargés $W^\pm$ et neutre Z^0, dont les masses sont respectivement 82 MeV/c^2 et 91 MeV/c^2, environ 100 fois la masse du proton. Les leptons, les quarks, les bosons de spin 1 (ou bosons de jauge) : photon, gluons, bosons $W^\pm$ et Z^0, ainsi qu'une particule hypothétique de spin 0, à l'origine de la masse des particules, le boson de Higgs, sont les particules du *modèle standard* de la physique des particules, qui a été testé expérimentalement avec une précision supérieure à 0.1 % au cours de ces dix dernières années.

Enfin les *interactions gravitationnelles*, toujours attractives contrairement aux interactions coulombiennes, ont la forme bien connue entre deux masses m et m'

$$\vec{F} = -Gmm' \, \frac{\hat{r}}{r^2} \tag{1.5}$$

où les notations sont identiques à celles de (1.2) et où G est la constante de gravitation. La loi de force (1.5), est, comme la loi de Coulomb, une loi à longue portée, et on peut comparer directement les forces de gravitation et de Coulomb entre un électron et un proton, puisque la forme de la loi de force est la même

$$\frac{F_{\mathrm{Cb}}}{F_{\mathrm{gr.}}} = \left(\frac{q_e^2}{4\pi\varepsilon_0}\right) \left(\frac{1}{Gm_e m_p}\right) \sim 10^{39}$$

Dans un atome d'hydrogène, la force de gravitation est négligeable, et de façon générale, la force de gravitation sera totalement négligeable pour tous les phénomènes de physique atomique, moléculaire ou du solide. La relativité générale, théorie relativiste de la gravitation, prédit l'existence d'ondes gravitationnelles[14] qui sont pour la gravitation l'analogue des ondes électromagnétiques, tandis que le graviton, particule de spin 2 et de masse nulle, est l'analogue du photon. Toutefois il n'existe pas aujourd'hui de théorie quantique de la gravitation. Concilier la mécanique quantique et la relativité générale, expliquer l'origine de la masse et des trois familles de particules constituent les défis majeurs de la physique théorique du XXIe siècle.

14. Les preuves de l'existence des ondes gravitationnelles sont pour le moment indirectes : elles résultent de l'observation de pulsars (étoiles à neutrons) binaires. Ces ondes pourraient être détectées prochainement sur Terre dans les expériences VIRGO et LIGO. L'observation du graviton est probablement repoussée à un futur très lointain.

Comme résumé de cette présentation des constituants élémentaires et des forces, on retiendra qu'il existe trois familles de particules de matière, comprenant des leptons et des quarks, et que les vecteurs des forces sont le photon pour les interactions électromagnétiques, le gluon pour les interaction fortes, les bosons W et Z pour les interactions faibles, et enfin l'hypothétique graviton pour les interactions gravitationnelles.

1.2 Physique classique et physique quantique

Avant d'introduire la physique quantique, résumons brièvement les fondements de la physique classique. La physique classique comporte trois branches principales, qui ont chacune diverses ramifications.

1. La première branche est la *mécanique*, dont la loi fondamentale est la loi de Newton, ou loi fondamentale de la dynamique, donnant la force $\vec{F}$ sur une particule ponctuelle de masse m en fonction de la dérivée par rapport au temps de son impulsion $\vec{p}$

$$\vec{F} = \frac{\mathrm{d}\vec{p}}{\mathrm{d}t} \tag{1.6}$$

Sous cette forme, l'équation fondamentale de la dynamique survit aux modifications apportées par Einstein en 1905 avec la relativité restreinte, à condition d'utiliser l'expression relativiste de l'impulsion en fonction de la vitesse $\vec{v}$, de la masse m de la particule et de la vitesse de la lumière c

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - v^2/c^2}} \tag{1.7}$$

2. La deuxième branche est *l'électromagnétisme*, résumé dans les quatre équations de Maxwell donnant le champ électrique $\vec{E}$ et le champ magnétique $\vec{B}$ en fonction des densités de charge ρ_{em} et de courant $\vec{\jmath}_{\mathrm{em}}$, appelées *sources* du champ électromagnétique

$$\vec{\nabla} \cdot \vec{B} = 0 \qquad \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{1.8}$$

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho_{\mathrm{em}}}{\varepsilon_0} \qquad c^2 \vec{\nabla} \times \vec{B} = \frac{\partial \vec{E}}{\partial t} + \frac{1}{\varepsilon_0}\,\vec{\jmath}_{\mathrm{em}} \tag{1.9}$$

De ces équations, on déduit la propagation d'ondes électromagnétiques dans le vide à la vitesse de la lumière

$$\left(\frac{1}{c^2}\,\frac{\partial^2}{\partial t^2} - \nabla^2\right)\left\{\begin{matrix}\vec{E}\\ \vec{B}\end{matrix}\right. = 0 \tag{1.10}$$

Ceci fait le lien avec l'optique, qui devient un cas particulier de l'électromagnétisme. Le lien entre **1** et **2** est fourni par la *loi de Lorentz* donnant la force sur une particule de charge q et de vitesse $\vec{v}$

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) \tag{1.11}$$

3. La troisième branche est la *thermodynamique*, qui se déduit du second principe[15] : il n'existe pas de dispositif fonctionnant suivant un cycle dont le seul effet serait de fournir du travail à partir d'une source de chaleur unique. À partir du second principe, on déduit la notion d'entropie, dont découle toute la thermodynamique classique. L'origine microscopique du second principe a été comprise à la fin du XIXe siècle par Boltzmann et Gibbs, qui ont pu relier ce principe au fait qu'un échantillon de matière macroscopique est composé d'un nombre énorme ($\sim 10^{23}$) d'atomes, ce qui permet d'utiliser des raisonnements probabilistes à la base de la mécanique statistique. Le résultat principal de la mécanique statistique est la *loi de Boltzmann* : la probabilité $\mathsf{p}(E)$ pour qu'un système physique en équilibre à la température absolue T ait l'énergie[16] E, comprend un facteur, le *poids de Boltzmann* $\mathsf{p}_{\rm B}(E)$

$$\mathsf{p}_{\rm B}(E) = \exp\left(-\frac{E}{k_{\rm B}T}\right) = \exp(-\beta E) \tag{1.12}$$

où $k_{\rm B}$ est la constante de Boltzmann (la constante R des gaz parfaits divisée par le nombre d'Avogadro), et nous avons introduit la notation usuelle $\beta = 1/(k_{\rm B}T)$. Cependant, la mécanique statistique classique n'est pas en fait une théorie cohérente, et il faut parfois se livrer à des acrobaties pour obtenir des résultats sensés, par exemple pour l'entropie d'un gaz parfait. La physique quantique lève toutes ces difficultés.

4. En toute rigueur il faut ajouter une quatrième branche à la physique classique. En effet, la théorie relativiste de la gravitation n'est pas incluse dans le cadre précédent : cette théorie est la *relativité générale*, qui est une description géométrique, où les forces de gravitation sont reliées à la courbure de l'espace-temps.

Nous avons décrit dans (1.6)–(1.11) les lois *fondamentales* de la physique classique, qui se résument donc à sept équations en tout et pour tout ! Le lecteur pourra se demander ce que sont devenues les multiples autres lois qu'il a rencontrées : loi d'Ohm, loi de Hooke, lois de la dynamique des fluides, etc. Certaines de ces lois se déduisent directement des lois fondamentales : ainsi la loi de Coulomb est une conséquence des équations de Maxwell et de la force de Lorentz (1.11) pour des charges statiques, l'équation d'Euler pour les fluides parfaits une conséquence de la loi fondamentale de la dynamique. D'autres lois sont des lois[17] *phénoménologiques*, qui n'ont pas une validité universelle, contrairement aux lois fondamentales : par exemple certains milieux n'obéissent pas à la loi d'Ohm, la relation entre l'induction $\vec{D}$ et le champ électrique $\vec{D} = \varepsilon\vec{E}$ (pour un milieu isotrope) est en défaut quand le champ

15. Le premier principe n'est autre que la conservation de l'énergie, quant au troisième, il est fondamentalement d'origine quantique.

16. La probabilité $\mathsf{p}(E)$ est le produit de $\mathsf{p}_{\rm B}(E)$ (1.12) et d'un facteur $\mathcal{D}(E)$, la "densité des niveaux d'énergie", qui, en physique classique, s'obtient par une intégration sur l'espace de phase : voir la note 22. Le calcul quantique de la densité de niveaux est décrit au § 8.5.2.

17. Bien souvent une loi phénoménologique n'est pas autre chose que le premier terme d'un développement de Taylor !

électrique devient grand, ce qui donne lieu aux phénomènes de l'optique non-linéaire, la loi de Hooke n'est plus valable si la tension devient trop importante, etc. La mécanique du solide, l'élasticité ou la mécanique des fluides découlent de (1.6) et de diverses lois phénoménologiques, comme la loi qui relie en mécanique des fluides force, gradient de vitesse et viscosité. Il importe de bien faire la distinction entre le petit nombre de lois fondamentales et les multiples lois phénoménologiques que la physique classique utilise, faute de mieux, pour décrire la matière.

Bien que la physique classique soit d'une utilité indiscutable, elle n'en présente pas moins une lacune de taille : alors que la physique se veut une théorie de la matière, la physique classique est complètement incapable *d'expliquer* le comportement de la matière étant donné ses constituants et les forces entre ces constituants[18]. Elle ne peut pas prédire l'existence des atomes, car on ne peut pas construire une échelle de longueur avec les constantes de la physique classique : masses et charges du noyau et des électrons[19]. Elle n'explique pas pourquoi le Soleil brille, pourquoi la vapeur de sodium émet une lumière jaune, elle n'a rien à dire sur les propriétés chimiques des alcalins, sur le fait que le cuivre conduit l'électricité alors que le soufre est un isolant, etc. Lorsque le physicien classique a besoin d'une propriété de la matière, une résistance électrique, une chaleur spécifique, il n'a pas d'autre choix que de la mesurer expérimentalement. Au contraire, la mécanique quantique a la prétention d'expliquer le comportement de la matière à partir des constituants et des forces. Naturellement, des prédictions précises à partir des premiers principes ne sont possibles que pour les systèmes les plus simples, comme l'atome d'hydrogène ou celui d'hélium. La complexité des calculs ne permet pas par exemple de prédire la structure cristalline de l'argent à partir des données sur cet atome, mais étant donné cette structure, on saura expliquer pourquoi l'argent est un conducteur, ce que la physique classique est incapable de faire.

Il ne faudrait pas conclure de cette discussion que la physique classique ne peut plus être intéressante et novatrice. Bien au contraire, on a assisté au cours de ces trente dernières années à un renouveau de la physique classique, avec des idées nouvelles sur les systèmes dynamiques chaotiques, les instabilités, les formations de structures hors équilibre, etc. Des problèmes aussi familiers

18. Cette affirmation mérite d'être un peu nuancée. Il existe de bons modèles microscopiques en physique classique : par exemple la théorie cinétique des gaz permet un calcul fiable des coefficients de transport (viscosité, conductibilité thermique) d'un gaz. Mais ni l'existence des molécules qui composent le gaz, ni la valeur de la section efficace nécessaire au calcul ne peuvent s'expliquer par la physique classique.

19. Si l'on ajoute la vitesse de la lumière, on peut construire une échelle de longueur, le *rayon classique de l'électron*

$$r_e = \frac{q_e^2}{4\pi\varepsilon_0}\,\frac{1}{m_e c^2} \simeq 2.8 \times 10^{-15}\,\mathrm{m}$$

qui est trop petit par 4 ordres de grandeur par rapport aux dimensions atomiques. On peut aussi invoquer l'invariance d'échelle des équations classiques : *cf.* Wichman [1974], chapitre 1.

que la turbulence ou le frottement restent largement ouverts et passionnants. Simplement, il existe des problèmes que par nature la physique classique ne *peut pas* aborder.

La physique quantique prétend donc *expliquer* le comportement de la matière à partir des constituants et des forces, mais il y a un prix à payer : les objets quantiques exhibent un comportement radicalement nouveau, qui défie notre intuition fondée sur l'expérience du comportement d'objets classiques. Cela dit, si l'on ne se préoccupe pas de cet aspect surprenant, la mécanique quantique se révèle un outil remarquable, jamais mis en défaut jusqu'à aujourd'hui, capable de couvrir des phénomènes allant de la physique des quarks à la cosmologie en passant par toutes les échelles intermédiaires. La plupart des technologies modernes n'auraient pas vu le jour sans la mécanique quantique : toutes les technologies de l'information sont fondées sur notre compréhension quantique des solides, en particulier des semi-conducteurs, et sur celle des lasers. On peut prévoir que la miniaturisation des dispositifs électroniques rendra la mécanique quantique de plus en plus omniprésente dans la technologie moderne.

La grande majorité des physiciens ne se préoccupent pas des propriétés déroutantes de la mécanique quantique et s'en servent comme d'un outil sans se poser de questions de principe. Cependant, les progrès théoriques et surtout expérimentaux de ces vingt dernières années ont permis de mieux cerner certains aspects du comportement des objets quantiques. Des expériences permettant de tester directement les fondements de la mécanique quantique, autrefois qualifiées "d'expériences de pensée" (gedanken experiment) par les pères fondateurs qui les jugeaient irréalisables[20], sont maintenant monnaie courante dans les laboratoires. En dépit de ces progrès spectaculaires que nous examinerons en particulier au chapitre 11, nombre de questions fondamentales restent ouvertes, et l'affirmation provocatrice de Feynman : "je pense que l'on peut dire aujourd'hui que personne ne comprend la mécanique quantique" garde encore aujourd'hui une part de vérité. Avant d'aborder ces développements récents, revenons quelques instants une centaine d'années en arrière, aux débuts de la physique quantique.

1.3 Un peu d'histoire

1.3.1 Le rayonnement du corps noir

Un objet chaud comme un fer chauffé au rouge, ou le Soleil, émet un rayonnement électromagnétique avec un spectre de fréquences qui dépend de la température. La puissance émise $u(\omega, T)$ par unité de fréquence ω et

20. Ainsi Schrödinger écrivait en 1952 : *Nous n'expérimentons jamais juste avec un seul électron ou une seule (petite) molécule. Dans les expériences de pensée, nous supposons parfois que nous le faisons : ceci entraîne invariablement des conséquences ridicules... On peut dire à juste titre que nous n'expérimentons pas plus avec des particules isolées que nous n'élevons d'ichtyosaures dans un zoo* (Schrödinger [1952]).

de surface dépend de la température absolue T de l'objet. Un raisonnement purement thermodynamique permet de montrer que si l'objet est parfaitement absorbant, ce qui en fait un *corps noir*, alors $u(\omega, T)$ est une fonction universelle, indépendante de l'objet pour une température donnée. Une très bonne réalisation d'un corps noir pour la lumière visible est une petite ouverture dans une enceinte dont l'intérieur est peint en noir : un rayon lumineux pénétrant dans l'enceinte n'a pratiquement aucune chance d'en ressortir, puisqu'à chaque réflexion il a une bonne probabilité d'être absorbé par la paroi intérieure de l'enceinte (figure 1.4).

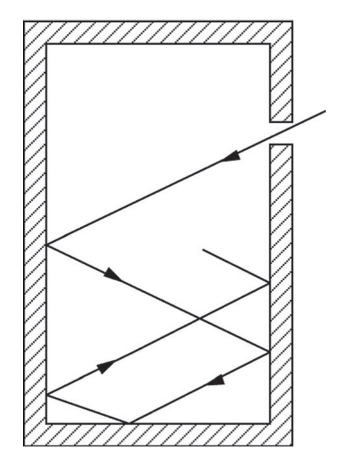


FIG. 1.4 – Enceinte pour le rayonnement du corps noir.

Supposons l'enceinte chauffée à la température T : les atomes de la paroi émettent et absorbent du rayonnement électromagnétique, et il s'établit à l'équilibre thermodynamique un système d'ondes stationnaires dans la cavité. Si la cavité est un parallélépipède de côtés L_x, L_y et L_z, et si nous utilisons des conditions aux limites périodiques, le champ électrique sera de la forme $\vec{E}_0 \exp[\mathrm{i}(\vec{k}\cdot\vec{r}-\omega t)]$, le vecteur d'onde $\vec{k}$ perpendiculaire à $\vec{E}_0$ étant de la forme

$$\vec{k} = \left(\frac{2\pi}{L_x}\, n_x,\ \frac{2\pi}{L_y}\, n_y,\ \frac{2\pi}{L_z}\, n_z\right) \tag{1.13}$$

où $(n_x,\ n_y,\ n_z)$ sont des nombres entiers positifs ou négatifs et $\omega = c|\vec{k}| = ck$. On montre que chaque onde stationnaire se comporte comme un oscillateur harmonique[21] de fréquence ω dont l'énergie est proportionnelle au carré de l'amplitude, et donc à $\vec{E}_0^2$. D'après la loi de Boltzmann (1.12), la probabilité que cet oscillateur ait une énergie E comprend un facteur $\exp(-E/k_\mathrm{B}T) = \exp(-\beta E)$. En fait, dans ce cas, la densité de niveaux $\mathcal{D}(E)$ (*cf.* note 16) est

21. Ceci sera expliqué au § 17.1.1.

une constante[22], et l'énergie moyenne de cet oscillateur est simplement

$$\begin{aligned} \langle E \rangle &= \frac{\int \mathrm{d}E\, E \exp(-\beta E)}{\int \mathrm{d}E\, \exp(-\beta E)} = -\frac{\partial}{\partial \beta} \ln \left(\int \mathrm{d}E\, \exp(-\beta E) \right) \\ &= -\frac{\partial}{\partial \beta} \ln \frac{1}{\beta} = \frac{1}{\beta} = k_{\mathrm{B}} T \end{aligned} \tag{1.14}$$

L'énergie moyenne de chaque onde stationnaire est $k_{\mathrm{B}}T$. Comme le nombre d'ondes stationnaires possibles est infini, l'énergie dans l'enceinte est infinie !

On relie facilement $u(\omega, T)$ à la densité d'énergie $\epsilon(\omega, T)$ par unité de fréquence dans l'enceinte (exercice 1.6.2)

$$u(\omega, T) = \frac{c}{4}\, \epsilon(\omega, T) \tag{1.15}$$

et on est ramené au calcul de $\epsilon(\omega, T)$, dont on déduit la densité d'énergie

$$\epsilon(T) = \int_0^\infty \mathrm{d}\omega\, \epsilon(\omega, T) \tag{1.16}$$

La thermodynamique permet de montrer la loi d'échelle

$$\epsilon(\omega, T) = \omega^3 \varphi \left(\frac{\omega}{T} \right) \tag{1.17}$$

où la fonction φ est indépendante de la forme de l'enceinte, mais elle ne dit rien sur la forme explicite de la fonction φ. Essayons de la déterminer à un facteur multiplicatif près par analyse dimensionnelle : *a priori*, $\epsilon(\omega, T)$ ne peut dépendre que de ω, de c, de l'énergie disponible dans le problème $k_{\mathrm{B}}T$ et d'une constante sans dimension A que l'analyse dimensionnelle ne permet pas de fixer. La seule solution possible est (exercice 1.6.2)

$$\epsilon(\omega, T) = A c^{-3} (k_{\mathrm{B}} T) \omega^2 = \omega^3 \left[A c^{-3} \left(\frac{k_{\mathrm{B}} T)}{\omega} \right) \right] \tag{1.18}$$

ce qui est bien de la forme (1.17). On retrouve le fait que la densité d'énergie dans l'enceinte est infinie

$$\epsilon(T) = \int_0^\infty \mathrm{d}\omega\, \epsilon(\omega, T) = A c^{-3} (k_{\mathrm{B}} T) \int_0^\infty \omega^2 \mathrm{d}\omega = +\infty$$

La mécanique statistique permet de calculer A (exercice 1.6.2), mais ne résout en rien le problème de l'énergie infinie, et l'analyse dimensionnelle suggère

22. L'intégration sur l'espace de phase pour l'oscillateur harmonique classique à une dimension donne en effet, pour une fonction arbitraire $f(E)$ (exercice 1.6.2)

$$\int \mathrm{d}\, x \mathrm{d}p\, \delta \left(E - \frac{p^2}{2m} - \frac{1}{2}\, m\omega^2 x^2 \right) f(E) = \frac{2\pi}{\omega}\, f(E)$$

x et p étant la position et l'impulsion.

fortement que le rayonnement du corps noir ne peut s'expliquer que si l'on accepte d'introduire une nouvelle constante physique.

Parmi toutes les hypothèses conduisant au résultat inacceptable d'une énergie infinie, Planck remet en cause celle qui conduit au calcul (1.14) de l'énergie moyenne d'un oscillateur[23] : au lieu de supposer que E peut prendre toutes les valeurs possibles entre zéro et l'infini, il admet que cette énergie ne peut prendre que des valeurs discrètes E_n qui sont des multiples entiers de la fréquence ω de l'oscillateur, avec un coefficient de proportionnalité $\hbar$

$$E_n = n(\hbar\omega) \qquad n = 0, 1, 2, \ \ldots \tag{1.19}$$

La constante $\hbar$ est appelée *constante de Planck*; plus exactement, c'est la constante de Planck h divisée[24] par 2π : $\hbar = h/(2\pi)$. La constante de Planck se mesure en J.s : elle a pour dimension $\mathcal{M}\mathcal{L}^2\mathcal{T}^{-1}$ et elle a pour valeur numérique

$$\hbar \simeq 1.055 \times 10^{-34}\,\text{J.s} \qquad \text{ou} \qquad h \simeq 6.63 \times 10^{-34}\,\text{J.s}$$

D'après la loi de Boltzmann, la probabilité normalisée d'observer une énergie E_n est

$$\mathsf{p}(E_n) = \mathrm{e}^{-\beta n\hbar\omega}\left(\sum_{n=0}^{\infty}\mathrm{e}^{-\beta n\hbar\omega}\right)^{-1} = \exp(-\beta n\hbar\omega)(1 - \exp(-\beta\hbar\omega)) \tag{1.20}$$

Pour obtenir (1.20), on remarque que la sommation sur n est celle d'une série géométrique. Posant $x = \exp(-\beta\hbar\omega)$, on calcule aisément la valeur moyenne $\langle E\rangle$ de l'énergie d'un oscillateur

$$\begin{aligned}\langle E\rangle &= (1-x)\sum_{n=0}^{\infty}(n\hbar\omega)x^n = (1-x)\hbar\omega x\frac{\mathrm{d}}{\mathrm{d}x}\sum_{n=0}^{\infty}x^n \\ &= (1-x)\hbar\omega x\frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{1-x} = \frac{\hbar\omega x}{1-x} = \frac{\hbar\omega}{\exp(\beta\hbar\omega)-1}\end{aligned} \tag{1.21}$$

Cette formule permet de calculer la densité d'énergie (exercice 1.6.2)

$$\epsilon(\omega, T) = \frac{\hbar}{\pi^2 c^3}\,\frac{\omega^3}{\exp(\beta\hbar\omega)-1} \tag{1.22}$$

23. En réalité, Planck a appliqué son raisonnement à un “résonateur” dont la nature reste obscure. Considérer les vibrations du champ électromagnétique est plus simple et plus direct, c'est ce que fait Einstein en 1905, mais constitue une entorse à la vérité historique. Notre présentation “historique”, tout comme celle de la plupart des manuels, tient plus du conte de fées (Kragh [2000]) que de l'histoire réelle. De même, il ne semble pas que les physiciens de la fin du dix-neuvième siècle aient été préoccupés par le problème de l'énergie infinie, ou par l'absence d'une constante fondamentale.

24. Nous utiliserons systématiquement $\hbar$ et non h, et par abus de langage nous appellerons $\hbar$ la constante de Planck; la relation $E = \hbar\omega$ est bien sûr équivalente à $E = h\nu$, où ν est la fréquence ordinaire, mesurée en Hz, et ω la fréquence angulaire, ou pulsation, mesurée en rad.s^{-1} : $\omega = 2\pi\nu$. Comme nous utiliserons pratiquement toujours ω et jamais ν, par abus de langage nous appellerons ω la fréquence.

et donc $u(\omega, T)$, en parfait accord avec l'expérience si l'on fixe convenablement la valeur de $\hbar$, et avec le résultat (1.17) de la thermodynamique. On remarque que l'approximation classique (1.18) est valable si $k_{\mathrm{B}}T \gg \hbar\omega$, c'est-à-dire pour les basses fréquences.

L'exemple le plus connu de rayonnement du corps noir est le rayonnement fossile qui remplit l'Univers[25], ou rayonnement à 3 K. La distribution de fréquence de ce rayonnement suit remarquablement la loi de Planck (1.22) avec une température de 2.73 K $\simeq$ 3 K (figure 1.5), mais ce rayonnement n'est plus à l'équilibre thermodynamique. Il s'est découplé des atomes environ 380 000 ans après le big-bang, c'est-à-dire la naissance de l'Univers. Au moment de ce découplage, le température était de 10^4 K environ. L'expansion de l'Univers a aujourd'hui réduit cette valeur à 3 K.

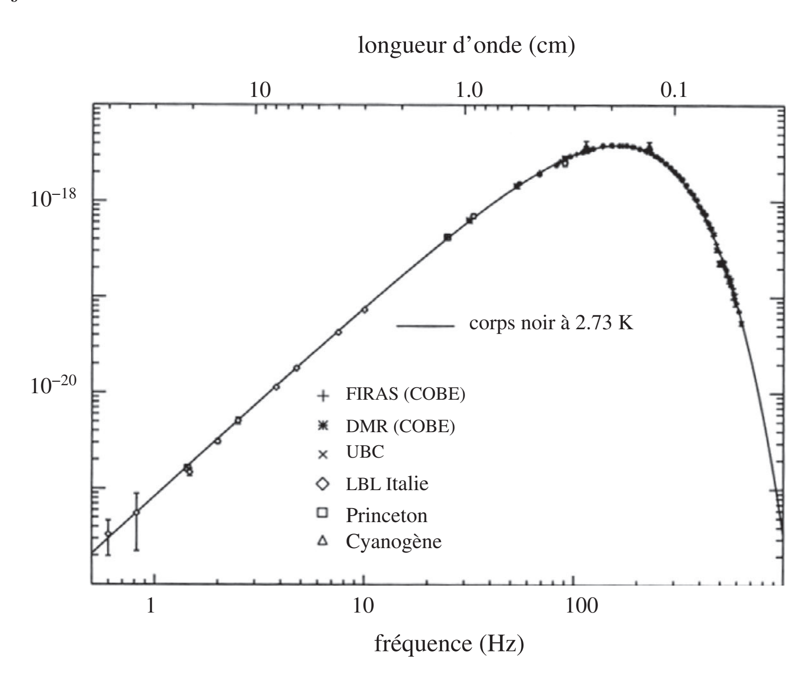


FIG. 1.5 – Le rayonnement du corps noir à 3 K. L'axe vertical donne l'intensité du rayonnement en W.m^{-2}.sr^{-1}.Hz^{-1}. On observera l'accord remarquable avec la loi de Planck pour $T = 2.73$ K. D'après J. Rich [2002].

1.3.2 L'effet photoélectrique

Le nombre entier n dans (1.19) possède une interprétation physique particulièrement importante : la raison pour laquelle l'énergie d'une onde stationnaire de fréquence ω est un multiple entier $n\hbar\omega$ de $\hbar\omega$ est que l'on y trouve

25. On trouvera un exposé remarquable du big bang dans Weinberg [1978].

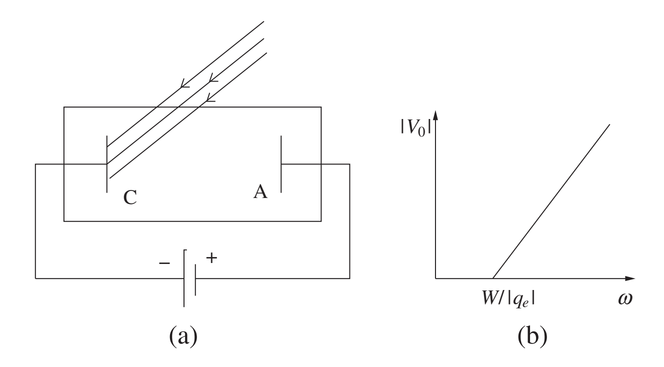


FIG. 1.6 – L'expérience de Millikan. (a) Schéma de l'expérience. (b) $|V_0|$ en fonction de ω.

précisément *n photons* (ou particules de lumière) d'énergie $\hbar\omega$. C'est cette interprétation qui a conduit Einstein à introduire le concept de photon pour expliquer l'effet photoélectrique. Lorsqu'un métal est illuminé par un rayonnement électromagnétique, des électrons sont arrachés au métal, avec un effet de seuil qui dépend de la fréquence, et non de l'intensité. L'expérience de Millikan (figure 1.6) confirme l'interpétation d'Einstein : les électrons sont arrachés au métal avec une énergie cinétique E_c

$$E_c = \hbar\omega - W \tag{1.23}$$

où W est le potentiel d'extraction. Aucun électron de charge q_e n'atteint la cathode si $|q_e V| > E_c$. Si V_0 est le potentiel pour lequel le courant s'annule

$$|V_0| = \frac{\hbar}{|q_e|}\,\omega - \frac{W}{|q_e|} \tag{1.24}$$

Porter $|V_0|$ en fonction de ω donne une droite de pente $\hbar/|q_e|$, et la valeur de $\hbar$ coïncide avec celle du rayonnement du corps noir, ce qui confirme l'hypothèse d'Einstein[26] : le rayonnement électromagnétique est composé de photons[27].

26. Encore une récriture de l'histoire ! Certains résultats qualitatifs sur l'effet photoélectrique avaient été obtenus par Lenard au début des années 1900, mais les mesures précises de Millikan sont postérieures de 10 ans à l'hypothèse d'Einstein, qui semble avoir été motivé non par l'effet photoélectrique, mais par des arguments thermodynamiques : voir par exemple Darrigol [2005].

27. Toutefois l'argument n'est pas entièrement convaincant, car l'effet photoélectrique peut s'expliquer dans le cadre d'une théorie semi-classique, où le champ électromagnétique n'est pas quantifié et où le concept de photon n'existe pas : *cf.* § 15.3.3. En revanche, on ne peut pas expliquer l'effet photoélectrique sans introduire $\hbar$. Le fait qu'un photomultiplicateur, dont le fonctionnement repose sur l'effet photoélectrique, enregistre des coups isolés peut être attribué au caractère quantique du détecteur et non à l'arrivée de photons isolés, voir le § 17.3.3.

1.4 Ondes et particules : interférences

1.4.1 Hypothèse de de Broglie

Partons de la relation (1.19) $E = \hbar\omega$ pour $n = 1$ reliant l'énergie et la fréquence d'un photon, aussi appelée *relation de Planck-Einstein*. Un photon possède une impulsion

$$p = \frac{E}{c} = \frac{\hbar\omega}{c}$$

mais compte tenu de $\omega = ck$ et de ce que l'impulsion et le vecteur d'onde sont parallèles et de même sens, on aboutit à la relation vectorielle suivante entre impulsion $\vec{p}$ et vecteur d'onde $\vec{k}$

$$\boxed{\vec{p} = \hbar\vec{k}} \tag{1.25}$$

Cette équation se traduit aussi par une relation (cette fois scalaire) entre l'impulsion et la longueur d'onde λ, la *longueur d'onde de de Broglie*

$$\boxed{p = \frac{h}{\lambda}} \tag{1.26}$$

L'hypothèse de de Broglie est que les relations (1.25) et (1.26) *sont valables pour toutes les particules*. Selon cette hypothèse, *une particule d'impulsion $\vec{p}$ possède des propriétés ondulatoires caractéristiques d'une longueur d'onde $\lambda = h/p$*. Si $v \ll c$, on utilisera $\vec{p} = m\vec{v}$, et sinon la formule générale (1.7), sauf bien sûr pour $m = 0$, où $p = E/c$. Si cette hypothèse est correcte, on doit pouvoir observer avec des particules des propriétés caractéristiques des ondes comme les interférences et la diffraction.

1.4.2 Diffraction et interférences avec des neutrons froids

Depuis les années 1980, les techniques expérimentales modernes permettent de vérifier les propriétés d'interférences et de diffraction de particules dans des expériences dont le principe est simple et dont l'interprétation est directe. Ces expériences ont été réalisées avec des photons, des électrons, des atomes, des molécules et des neutrons. Nous avons choisi, un peu arbitrairement, d'exposer les expériences réalisées avec des neutrons, qui nous ont semblé particulièrement élégantes et éclairantes. Les expériences de diffraction de neutrons par des cristaux sont classiques depuis plus de cinquante ans (exercice 1.6.4), mais l'idée est ici de réaliser des expériences avec des dispositifs macroscopiques, des fentes visibles à l'oeil nu, et non d'utiliser un réseau dont le pas est de quelques fractions de nm.

Les expériences ont été réalisées dans les années 1980 par un groupe d'Innsbruck auprès du réacteur nucléaire de recherche de l'Institut Laue-Langevin

à Grenoble. Les neutrons de masse m_n sont produits par la fission d'atomes d'uranium^{235} dans le cœur du réacteur, et sont ensuite guidés vers les expériences. En ordre de grandeur, leur énergie cinétique est $k_\mathrm{B}T$, où $T \sim 300\,\mathrm{K}$ est la température ambiante : on appelle ces neutrons des neutrons thermiques dont l'énergie cinétique $\sim k_\mathrm{B}T \simeq 1/40\,\mathrm{eV}$ pour $T = 300\,\mathrm{K}$. L'impulsion $p = \sqrt{2m_n k_\mathrm{B}T}$ correspond à une vitesse $v = p/m_n$ d'environ 1000 m.s^{-1} et d'après (1.26), la longueur d'onde λ_th vaut $h/\sqrt{2m_n k_\mathrm{B}T} \simeq 0.18\,\mathrm{nm}$. On augmente la longueur d'onde en faisant passer les neutrons dans des matériaux à basse température : par exemple si la température du matériau est 1 K, la longueur d'onde passera à $\lambda = \lambda_\mathrm{th}\sqrt{300} \simeq 3.1\,\mathrm{nm}$. De tels neutrons sont appelés "neutrons froids". Dans l'expérience du groupe d'Innsbruck, les neutrons sont "refroidis" dans du deutérium[28] liquide à 25 K. En sélectionnant les neutrons après leur passage dans le deutérium liquide, on obtient des neutrons dont la longueur d'onde moyenne est de 2 nm.

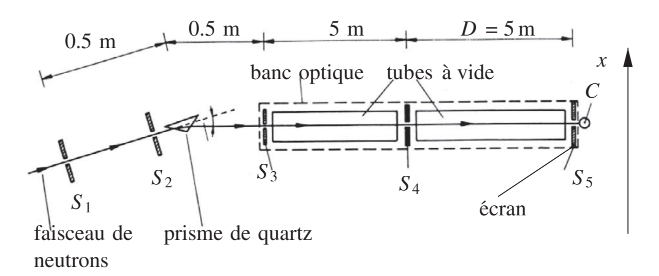


FIG. 1.7 – Dispositif expérimental pour la diffraction et les interférences de neutrons. S_1 et S_2 : fentes collimatrices. S_3 : fente d'entrée. S_4 : fente objet. S_5 : position du compteur C. D'après Zeilinger *et al.* [1988].

Le dispositif expérimental est schématisé sur la figure 1.7. La détection des neutrons se fait à l'aide de compteurs à fluorure de bore BF_3, le bore absorbant les neutrons suivant la réaction

$$^{10}\mathrm{B} + n \to \ ^7\mathrm{Li} + ^4\mathrm{He}$$

avec une efficacité voisine de 100 %. Le compteur est déplacé suivant l'écran en S_5, et compte le nombre de neutrons arrivant dans le voisinage de S_5.

Dans l'expérience de diffraction, la fente S_4 a une largeur $a = 93\,\mu\mathrm{m}$, ce qui donne une dimension angulaire de la tache de diffraction de

$$\theta = \frac{\lambda}{a} \sim 2 \times 10^{-5} \text{ radian}$$

et sur l'écran situé à $D = 5\,\mathrm{m}$ de la fente, une dimension linéaire de l'ordre de $100\,\mu\mathrm{m}$. Il est possible de faire un calcul précis de la figure de diffraction

28. Le deutérium est choisi de préférence à l'hydrogène, qui a l'inconvénient d'absorber les neutrons dans la réaction $n + p \to \ ^2\mathrm{H} + \gamma$. C'est pourquoi, dans un réacteur nucléaire, l'eau lourde est un meilleur modérateur que l'eau ordinaire : exercice 17.5.10.

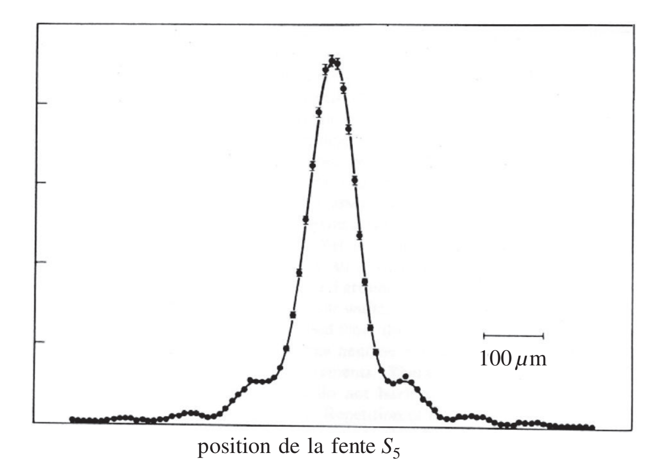


FIG. 1.8 – Diffraction de neutrons par une fente. D'après Zeilinger *et al.* [1988].

en tenant compte par exemple de la dispersion des longueurs d'onde autour de la longueur d'onde moyenne de 2 nm. Le résultat théorique est en accord remarquable avec l'expérience (figure 1.8).

Dans l'expérience d'interférences, deux fentes de 21 μm ont leurs centres espacés de $d = 125\,\mu$m. L'interfrange sur l'écran vaut

$$i = \frac{\lambda D}{d} = 80\ \mu\text{m}$$

Les fentes sont visibles à l'oeil nu, et l'interfrange est macroscopique. À nouveau, un calcul théorique prenant en compte les divers paramètres de l'expérience est en excellent accord avec la figure d'interférences expérimentale (figure 1.9).

Il y a toutefois une différence cruciale par rapport à une expérience d'interférences en optique : *la figure d'interférences est construite à partir d'impacts de neutrons isolés*, et elle est reconstituée après coup lorsque l'expérience est terminée. En effet, on déplace le compteur le long de l'écran (ou bien on dispose une batterie de compteurs identiques recouvrant l'écran), et on enregistre les neutrons arrivant au voisinage de chaque point de l'écran pendant des intervalles de temps identiques. Soit $N(x)\Delta x$ le nombre de neutrons détectés par seconde dans l'intervalle $[x - \Delta x/2, x + \Delta x/2]$, x étant l'abscisse d'un point sur l'écran. L'intensité $\mathcal{I}(x)$ peut être définie comme étant égale à $N(x)$ et le nombre de neutrons arrivant au voisinage d'un point de l'écran est proportionnel à l'intensité $\mathcal{I}(x)$ de la figure d'interférences, avec des fluctuations statistiques autour d'une valeur moyenne. Les impacts isolés sont illustrés sur la figure 1.10 par une expérience faite non avec des neutrons, mais des atomes froids que l'on laisse tomber à travers des fentes d'Young : les impacts

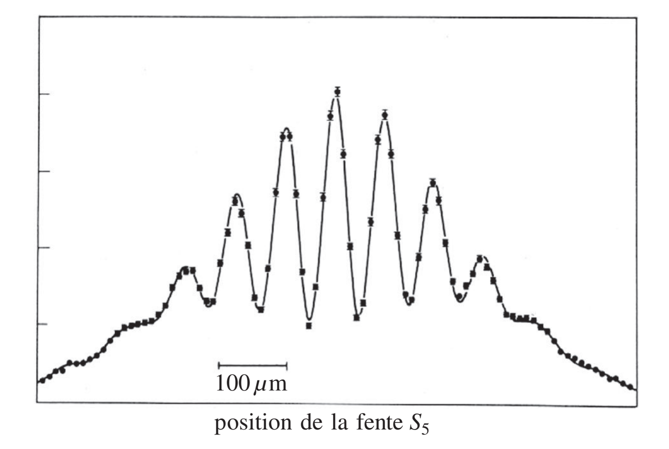


FIG. 1.9 – Expérience des fentes d'Young avec des neutrons. D'après Zeilinger *et al.* [1988].

des atomes tombant sur l'écran sont enregistrés pour donner l'aspect de la figure 1.10.

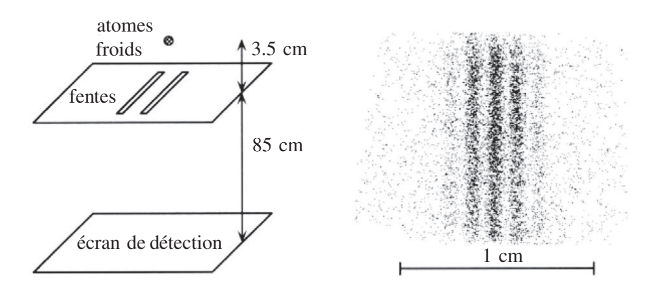


FIG. 1.10 – Interférences avec des atomes froids. D'après Shimizu *et al.* [1992].

1.4.3 Interprétation des expériences

Outre les neutrons et les atomes froids, les expériences de diffraction et d'interférences ont été réalisées avec plusieurs autres types de particules :

- avec des photons, en utilisant une source de photons uniques (Aspect *et al.* [1989], Jacques *et al.* [2005]). Les expériences effectuées en réduisant l'intensité lumineuse ne sont pas du tout probantes, car on peut en donner une explication semi-classique en prenant en compte le caractère quantique du détecteur : voir la note 27. Ce point délicat est examiné en détail au paragraphe suivant.
- avec des électrons

- avec des molécules légères (Na_2)
- avec des fullerènes C_{60} (exercice 1.6.1)

et il y a tout lieu de penser que les résultats sont universels, indépendants du type de particule : atomes, molécules, virus[29]... Cependant ces résultats expérimentaux incontournables semblent souffrir d'une difficulté d'interprétation : dans une expérience classique d'interférences de fentes d'Young réalisée avec des ondes, l'onde incidente se divise en deux ondes qui se recombinent ensuite et interfèrent, phénomène visible à l'oeil nu pour des ondes à la surface de l'eau. Dans le cas des neutrons, chaque neutron arrive isolément ; l'intervalle de temps entre deux neutrons successifs est tel que lorsqu'un neutron est détecté sur l'écran, le neutron suivant est encore dans le réacteur, confiné dans un atome d'uranium. Est-il envisageable que le neutron se scinde en deux fractions de neutron, chaque fraction passant par une fente ? Il est facile de se convaincre que cette hypothèse est absurde : un compteur détecte toujours un neutron entier, jamais une fraction de neutron !

Le cas des photons est plus complexe : on pourrait imaginer obtenir une source de photons uniques en atténuant une lumière classique, provenant par exemple d'une diode électroluminescente (LED) ou d'une diode laser. Prenons le cas d'une diode laser émettant dans le rouge une puissance de 1 mW. Connaissant la fréquence $\nu \simeq 3.8 \times 10^{14}$ Hz de la lumière émise, on déduit le flux de photons, environ 4×10^{15} photons par seconde. Atténuons le faisceau lumineux en réduisant sa puissance à 10^{-12} W. Dans ces conditions, la distance moyenne entre deux photons est de 100 m, ce qui est largement supérieur à la longueur de cohérence (voir § 5.4.2) de la lumière laser, et il semble raisonnable d'affirmer que l'on a construit une source de photons uniques. Cependant, cette affirmation est incorrecte. Pour le voir, utilisons le dispositif de la figure 1.11 : une *lame séparatrice*, ou *diviseur de faisceau*, divise le faisceau en un faisceau réfléchi vers le détecteur D_1 et un faisceau transmis vers le détecteur D_2. Nous utiliserons souvent des *lames séparatrices équilibrées* telles que 50 % de l'intensité incidente est réfléchie et 50 % transmise. Si l'on remplace la source lumineuse par une source de photons uniques, cela veut dire qu'un photon sera réfléchi une fois sur deux et transmis une fois sur deux, et un seul des deux détecteurs sera déclenché. Ce n'est pas le cas avec une source de lumière, laser ou LED. En effet, si $\overline{I}$ est l'intensité lumineuse moyenne incidente, l'intensité moyenne détectée en D_1, I_1 et en D_2, I_2, sera proportionnelle à $\overline{I}$, tandis que l'intensité recueillie par l'ensemble des détecteurs, I_{12} est proportionnelle à $\overline{I^2}$

$$I_1 = \eta \overline{I} \qquad I_2 = \eta \overline{I} \qquad I_{12} = \eta^2 \overline{I^2}$$

29. Cependant, l'observation d'effets ondulatoires est de plus en plus difficile quand les objets deviennent plus gros, d'abord parce que les longueurs d'onde sont de plus en plus courtes, mais aussi parce que les effets de la décohérence (§ 11.4.2) sont de plus en plus importants lorsque la taille de l'objet augmente : voir les articles de revue de Arndt *et al.* [2005], de Hornberger *et al.* [2012] et le § 17.3.5.

où η est un coefficient de proportionnalité. De la propriété générale $\overline{I^2} \geq (\overline{I})^2$, on déduit que $I_{12}/I_1 I_2 \geq 1$. Passant maintenant au cas des photons, on définit par analogie avec les intensités, les probabilités p_1 et p_2 de déclencher les détecteurs D_1 et D_2 et la probabilité p_{12} de déclenchement simultané, ainsi que le rapport

$$\alpha = \frac{\mathsf{p}_{12}}{\mathsf{p}_1 \mathsf{p}_2}$$

L'expérience (Jacques *et al.* [2005]) faite avec la lumière atténuée d'une LED donne $\alpha = 1.07 \pm 0.08$, ce qui s'explique par le fait que l'onde lumineuse est un état cohérent du champ (§ 17.2.1) : la lumière atténuée contient 0.01 photon par impulsion, et la probabilité de trouver deux photons dans l'impulsion est de 0.5×10^{-4}. *Une expérience d'interférences en lumière atténuée n'est pas une expérience faite avec des photons uniques.* Le critère de lumière non-classique est $\alpha < 1$, et avec des photons uniques, il faut en théorie atteindre $\alpha = 0$. En raison des imperfections expérimentales, il est impossible d'atteindre cette limite, mais on arrive tout de même à $\alpha = 0.12 \pm 0.02$ avec la source de photons uniques utilisée par Jacques *et al.* [2005].

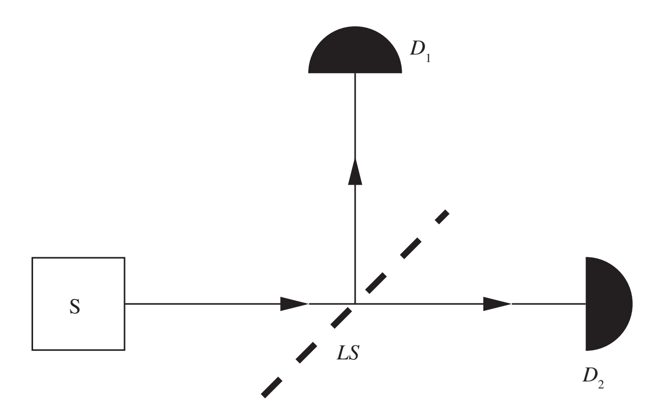


Fig. 1.11 – Lame séparatrice *LS* et comptage de photons par des photodétecteurs D_1 et D_2. La lame séparatrice réfléchit ou transmet un photon avec une probabilité de 50 %.

Il nous faut donc admettre qu'une particule quantique présente *à la fois* un aspect ondulatoire et un aspect corpusculaire : c'est donc *un objet entièrement nouveau et étrange*, du moins pour notre intuition formée par la pratique d'objets macroscopiques. Comme l'écrivent Lévy-Leblond et Balibar, paraphrasant Feynman “les objets quantiques sont complètement cinglés”, mais ils ajoutent “au moins le sont-ils tous de la même façon”. En effet photons, électrons, neutrons, atomes, molécules... ont tous ce même comportement, à la fois ondulatoire et corpusculaire parfois appelé “dualité onde-

corpuscule", une expression aujourd'hui obsolète (Kaiser et al. [2012]). Afin de mettre en valeur cette unicité du comportement quantique, certains auteurs ont proposé le néologisme "quanton" pour désigner un objet doté d'un tel comportement. Nous continuerons à utiliser "particule quantique", ou simplement "particule", car les particules considérées dans ce livre auront un comportement quantique, et nous préciserons "particule classique" si nous voulons revenir à des petites boules de billard.

Si le neutron est insécable, peut-on savoir s'il est passé par une fente plutôt que l'autre ? Si une fente est fermée, on observe sur l'écran la figure de diffraction de l'autre fente et réciproquement. Si la situation expérimentale est telle que l'on peut décider par quelle fente est passé le neutron, alors on doit observer sur l'écran la superposition des *intensités* des figures de diffraction de chaque fente : en effet on peut diviser les neutrons en deux groupes, ceux qui sont passés par la fente supérieure, et pour lesquels on aurait pu fermer la fente inférieure sans rien changer au résultat, et ceux qui sont passés par la fente inférieure. On n'observe une figure d'interférences que si le dispositif expérimental est tel que l'on ne peut pas savoir, *même en principe*, par quelle fente est passé le neutron. En résumé :

(i) Si le dispositif expérimental ne permet pas de savoir par quelle fente est passé le neutron, alors on observera des interférences.

(ii) Si le dispositif permet en principe de décider entre les deux fentes, alors les interférences seront détruites, indépendamment du fait qu'un expérimentateur se donne ou non la peine de faire l'observation nécessaire.

Une remarque fondamentale est que l'on ne peut pas savoir *a priori* en quel point de l'écran va arriver un neutron donné. On peut seulement affirmer que la *probabilité* d'arrivée sur l'écran est grande en un point de maximum de la figure d'interférences, et petite en un point où la figure d'interférences présente un minimum. Plus précisément, la probabilité d'arrivée en un point d'abscisse x est proportionnelle à l'intensité $\mathcal{I}(x)$ de la figure d'interférences en ce point. De même, dans l'expérience de la figure 1.11, chaque photomultiplicateur a une probabilité $1/2$ d'être déclenché par un photon donné, mais il est impossible de savoir à l'avance lequel des deux le sera.

Essayons de donner une formulation quantitative de la discussion précédente. Tout d'abord, par analogie avec les ondes, nous introduirons une fonction complexe de x, $a_1(x)$, (resp. $a_2(x)$) associée au passage par la fente supérieure (resp. inférieure) d'un neutron arrivant en x sur l'écran, appelée *amplitude de probabilité*. Le module au carré de l'amplitude de probabilité donne l'intensité : si la fente 2 est fermée $\mathcal{I}_1(x) \propto |a_1(x)|^2$ et inversement $\mathcal{I}_2(x) \propto |a_2(x)|^2$ si la fente 1 est fermée. Dans le cas, (i), on ajoute les *amplitudes* avant de calculer l'intensité

$$\mathcal{I}(x) \propto |a_1(x) + a_2(x)|^2 \tag{1.27}$$

et dans le cas (ii), on ajoute les *intensités*

$$\mathcal{I}(x) \propto |a_1(x)|^2 + |a_2(x)|^2 = \mathcal{I}_1(x) + \mathcal{I}_2(x) \tag{1.28}$$

Comme ci-dessus, l'intensité peut être définie comme le nombre de neutrons arrivant par seconde et par unité de longueur sur l'écran. Pour tenir compte du caractère probabiliste du point d'impact des neutrons, les amplitudes a_1 et a_2 ne seront pas des amplitudes ondulatoires, mesurant l'amplitude d'une vibration, mais des *amplitudes de probabilité*, dont le module carré donnera la *probabilité* d'arriver au point x sur l'écran. La notion d'amplitude de probabilité, propre à la physique quantique, sera développée et recevra un habillage mathématique au chapitre 3.

Une formulation plus générale de (1.27) et (1.28) est alors la suivante : supposons que partant d'un état initial i, on arrive à un état final f. Pour obtenir la probabilité $\mathsf{p}_{i\to f}$ d'observer l'état final f, on doit additionner toutes les amplitudes conduisant au résultat f en partant de i

$$a_{i\to f} = a^{(1)}_{i\to f} + a^{(2)}_{i\to f} + \cdots + a^{(n)}_{i\to f}$$

et $\mathsf{p}_{i\to f} = |a_{i\to f}|^2$. Il doit être bien entendu que les états i et f sont spécifiés de façon unique par la donnée des paramètres qui définissent l'état initial et l'état final de *l'ensemble* du dispositif expérimental. Si, par exemple, nous recherchons une information sur le passage du neutron à travers une fente, cette information ne peut être obtenue qu'en intégrant les fentes d'Young dans un dispositif plus vaste, dont l'état final, qui dépendra d'autres paramètres que le point d'impact du neutron, est susceptible de nous renseigner sur le passage par une fente déterminée : l'état final de l'ensemble du dispositif ne sera pas le même selon que le neutron est passé par une fente ou par l'autre. Dans le langage du chapitre 11, le passage du neutron par une fente plutôt que l'autre a laissé une trace dans l'environnement, ce qui conduit à la disparition des interférences.

En résumé, on doit sommer les *amplitudes* pour des états finaux[30] identiques, et les *probabilités* pour des états finaux différents, même si ces états finaux concernent d'autres paramètres physiques que ceux auxquels on s'intéresse. Il suffit que ces paramètres soient accessibles en principe, même s'ils ne sont pas effectivement observés, pour que l'on doive considérer des états finaux comme différents. Nous illustrerons ce point au paragraphe suivant sur un exemple concret. Une façon imagée d'exprimer les propriétés ci-dessus consiste à dire que tous les chemins conduisant à des états finaux identiques sont des *chemins indiscernables*, et que l'on doit sommer les amplitudes correspondant à tous ces chemins.

1.4.4 Inégalités de Heisenberg I

Revenons sur l'expérience de diffraction des neutrons pour en tirer une relation fondamentale, appelée *inégalité de Heisenberg*, ou suivant une terminologie courante mais ambiguë, *principe d'incertitude de Heisenberg*. Si la

30. La grammaire classique imposerait "états finals"...

largeur de la fente est a, et si nous orientons l'axe des x dans le plan de la fente perpendiculairement à celle-ci, la position du neutron suivant cet axe immédiatement à la sortie de la fente est précise à $\Delta x = a$ près. Comme la largeur angulaire de la tache de diffraction est $\sim \lambda/\Delta x$, la composante suivant x de l'impulsion du neutron est $\Delta p_x \sim (\lambda/\Delta x)p = h/\Delta x$, où p est l'impulsion du neutron (on suppose $p \gg \Delta p_x$). Nous avons donc la relation

$$\Delta p_x \, \Delta x \sim h \tag{1.29}$$

Nous verrons au chapitre 8 une version plus précise de (1.29), où Δp_i, $i = x, y, z$ et Δx_i, représenteront les écarts types, ou dispersions, sur des composantes identiques (i) de l'impulsion et de la position.

$$\Delta p_i \, \Delta x_i \geq \frac{1}{2} \, \hbar \tag{1.30}$$

En revanche, aucune inégalité ne relie des composantes différentes de l'impulsion et de la position : par exemple Δp_x et Δy ne sont contraints par aucune relation. On dit souvent, en interprétant l'expérience de diffraction, que le passage du neutron dans une fente de largeur Δx a permis de mesurer sa position suivant x avec une précision Δx, et que cette mesure a perturbé son impulsion par une quantité $\Delta p_x \sim h/\Delta x$. Nous verrons au chapitre 4 que les inégalités (1.30) n'ont en fait rien à voir avec une mesure expérimentale de la position ou de l'impulsion, mais proviennent de la description mathématique d'une particule quantique par un train d'ondes. Nous reviendrons également sur la signification des ces relations.

Nous allons maintenant utiliser (1.29) pour discuter la question de l'observation des trajectoires dans l'expérience d'interférences avec des neutrons. Einstein avait proposé le dispositif de la figure 1.12 pour déterminer la trajectoire du neutron : passe-t-il par la fente supérieure ou inférieure ? Quand le neutron franchit la première fente F_0, il donne, par conservation de l'impulsion, une impulsion vers le bas à l'écran E_0 s'il franchit la fente supérieure F_1 et une impulsion vers le haut s'il franchit la fente inférieure F_2. On peut donc déterminer par quelle fente est passé le neutron ! La réponse de Bohr fut la suivante : si l'écran E_0 reçoit une impulsion δp_x que l'on peut mesurer, alors cela veut dire que l'impulsion *initiale* Δp_x de l'écran était très inférieure à δp_x, et sa position initiale déterminée au mieux avec une précision de l'ordre de $h/\Delta p_x$. Cette imprécision dans la position de la source suffit à faire disparaître la figure d'interférences (exercice 1.6.3). Tous les dispositifs imaginés pour déterminer la trajectoire du neutron sont, soit efficaces, mais dans ce cas il n'y a pas d'interférences, soit inefficaces, et dans ce cas les interférences persistent, mais on ne sait pas par quelle fente est passé le neutron. La figure d'interférences se brouille au fur et à mesure que le dispositif devient de plus en plus efficace.

La discussion ci-dessus est en tout point correcte, mais elle masque le point essentiel : ce n'est pas la perturbation causée à la trajectoire du neu-

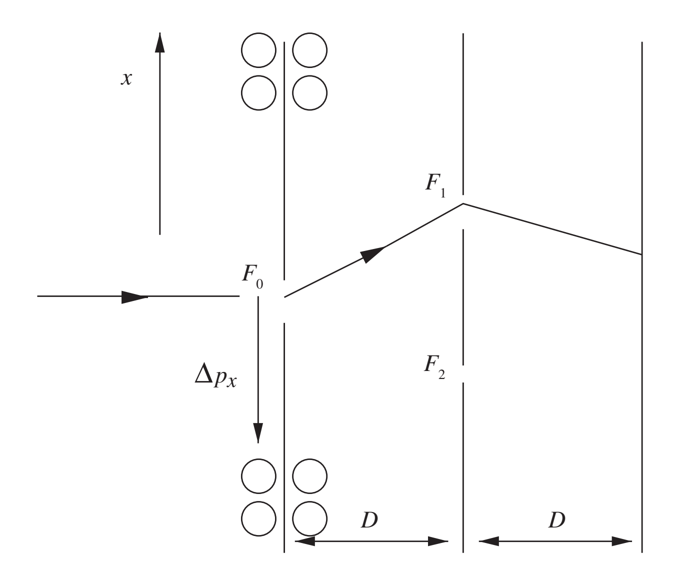


FIG. 1.12 – Une controverse Bohr-Einstein. Les fentes F_1 et F_2 sont les fentes d'Young. La fente F_0 est percée dans un écran mobile verticalement.

tron par le choc sur le premier écran qui brouille les interférences[31]. Ce qui est crucial est la possibilité *d'étiqueter la trajectoire*. On peut imaginer et même réaliser expérimentalement des dispositifs qui étiquettent les trajectoires sans perturber en quoi que ce soit les degrés de liberté observés, et cet étiquetage suffit à détruire les interférences. Nous allons décrire brièvement un tel dispositif, qui n'a pas encore été réalisé expérimentalement, mais qui ne semble pas hors de portée de développements technologiques futurs. Le dispositif proposé[32] utilise des atomes, ce qui permet de jouer sur leurs degrés de liberté internes sans affecter la trajectoire de leur centre de masse. Avant leur passage à travers les fentes d'Young, les atomes sont portés dans un état excité par un faisceau laser (figure 1.13). Derrière chacune des fentes d'Young se trouve une cavité supraconductrice micro-onde, décrite plus en détail au § 17.2.3. L'atome en passant dans la cavité revient à son état

31. La même remarque vaut pour le dispositif imaginé par Feynman pour une expérience de fentes d'Young avec des électrons (Feynman *et al.* [1965], vol. III, chapitre 1). Une source de photons placée derrière les fentes permet en théorie d'observer le passage des électrons. Lorsque l'on utilise des photons de courte longueur d'onde, les collisions électron-photon permettent de discriminer entre les fentes, mais les collisions perturbent suffisamment les trajectoires pour brouiller les interférences. Si on augmente la longueur d'onde, les chocs sont moins violents, mais le pouvoir de résolution des photons diminue. Les interférences réapparaissent quand ce pouvoir de résolution ne permet plus de distinguer entre les fentes. Voir le § 17.3.5 pour une variante de cette expérience qui a été effectivement réalisée.

32. Ce dispositif a été imaginé par Englert *et al.* [1991], et une version grand public en est donnée dans Englert *et al.* [1995]. Les atomes sont supposés se trouver dans des états de Rydberg (*cf.* exercice 15.4.4.). Une expérience voisine dans son principe, mais dont le schéma est plus complexe, a été effectivement réalisée par Dürr *et al.* [1998].

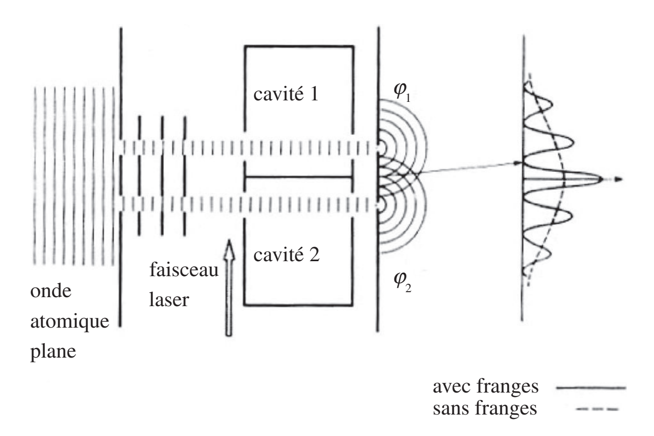


FIG. 1.13 – Étiquetage des trajectoires dans l'expériences des fentes d'Young. D'après Englert *et al.* [1991].

fondamental en émettant avec une probabilité voisine de 100 % un photon qui reste confiné dans la cavité. La présence du photon dans l'une ou l'autre des cavités permet d'étiqueter la trajectoire de l'atome, ce qui détruit les interférences. La perturbation apportée à la trajectoire du centre de masse de l'atome est totalement négligeable : il n'y a pratiquement aucun transfert d'impulsion entre le photon et l'atome. Cependant, les deux états finaux : atome arrivant au point d'abscisse x sur l'écran et photon dans la cavité 1, et atome arrivant au point x sur l'écran et photon dans la cavité 2 sont *différents*. Il faut donc prendre le module carré de chacune des amplitudes correspondantes et ajouter les probabilités. On note qu'il n'est pas nécessaire de détecter le photon, ce qui introduirait d'ailleurs une complication expérimentale supplémentaire. Il suffit de savoir que l'atome émet un photon de façon quasi-certaine au cours de son passage dans la cavité. Comme nous l'avons déjà souligné, il n'est pas indispensable que l'observation de l'état final soit effectivement réalisée et il suffit qu'elle soit possible en principe : peu importe que la technologie d'aujourd'hui ou même de demain soit capable ou non de permettre cette observation.

1.4.5 Interféromètre de Mach-Zehnder

Les fentes d'Young sont l'exemple standard d'une expérience d'interférences, mais un interféromètre beaucoup plus largement utilisé pour réaliser des expériences de physique quantique, avec des photons, des neutrons, des atomes, etc., est celui de Mach-Zehnder, qui repose sur l'utilisation de lames séparatrices (figure 1.14a). Nous allons partir du point de vue ondulatoire en nous plaçant dans le cadre de l'optique. Le schéma de l'interféromètre est donné sur la figure 1.14a. Examinons l'action d'une lame séparatrice sur des

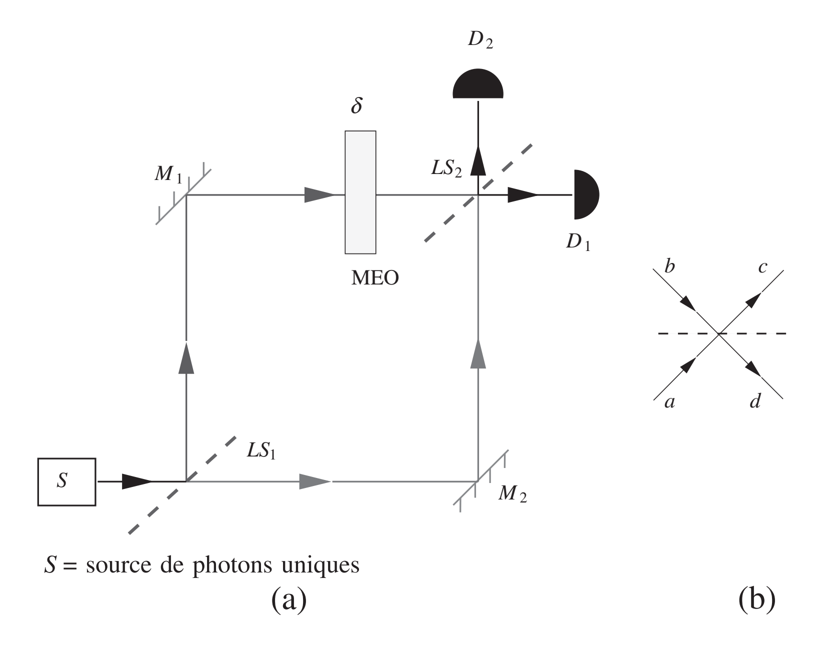


FIG. 1.14 – (a) Interféromètre de Mach-Zehnder : un photon unique est incident sur la lame séparatrice LS_1. Les deux miroirs M_1 et M_2 recombinent les amplitudes sur la lame séparatrice LS_2 et le photon est détecté en D_1 ou D_2. Un modulateur électro-optique MEO introduit un déphasage δ. (b) Action d'une lame séparatrice.

amplitudes ondulatoires : soit $\overline{a}$ et $\overline{b}$ les amplitudes d'une onde arrivant de chaque côté de la lame séparatrice, et soit $\overline{c}$ et $\overline{d}$ les amplitudes ondulatoires quittant la lame séparatrice (figure 1.14b). Dans le cas des ondes électromagnétiques, ces amplitudes peuvent être identifiées par exemple aux champs électriques. En raison du principe de superposition, les amplitudes $\overline{c}$ et $\overline{d}$ sont reliées linéairement aux amplitudes $\overline{a}$ et $\overline{b}$

$$\begin{aligned}\overline{c} &= t\overline{a} + r\overline{b} \\ \overline{d} &= r\overline{a} + t\overline{b}\end{aligned} \tag{1.31}$$

Dans cette équation, t décrit la transmission et r la réflexion. Nous avons supposé la lame séparatrice *symétrique* : les coefficients t et r sont identiques pour les deux lignes de (1.31), ce qui peut se déduire du principe de retour inverse de la lumière ; pour le cas non symétrique, voir Zeilinger [1981]. L'équation (1.31) s'écrit sous forme matricielle

$$\begin{pmatrix}\overline{c} \\ \overline{d}\end{pmatrix} = \begin{pmatrix} t & r \\ r & t \end{pmatrix} \begin{pmatrix}\overline{a} \\ \overline{b}\end{pmatrix} = S \begin{pmatrix}\overline{a} \\ \overline{b}\end{pmatrix} \tag{1.32}$$

Nous supposons également la lame séparatrice sans pertes : l'intensité lumineuse à la sortie doit être égale à l'intensité d'entrée

$$|\overline{c}|^2 + |\overline{d}|^2 = |t\overline{a} + r\overline{b}|^2 + |r\overline{a} + t\overline{b}|^2 = |\overline{a}|^2 + |\overline{b}|^2 \tag{1.33}$$

Cette condition est équivalente à l'unitarité de la matrice S : $S^\dagger S = I$. On déduit de (1.33) les conditions suivantes sur r et t

$$|t|^2 + |r|^2 = 1 \qquad \mathrm{Re}(t^* r) = 0 \tag{1.34}$$

Si les phases α et β de t et r sont définies par $t = |t| \exp(\mathrm{i}\alpha)$ et $r = |r| \exp(\mathrm{i}\beta)$, l'équation(1.34) implique

$$\alpha - \beta = \frac{\pi}{2} \pm n\pi \qquad n = 0, 1, 2, \ \ldots \tag{1.35}$$

Comme la phase globale de t et r est sans pertinence physique, on peut choisir t réel positif, $t = |t|$, et $r = \pm\mathrm{i}|r|$.

Le raisonnement ci-dessus se transpose immédiatement au cas de particules quantiques : il suffit de remplacer les amplitudes ondulatoires $\overline{a}, \cdots, \overline{d}$ par des amplitudes de probabilité $a, \cdots, d$. Dans le schéma de la figure 1.14, l'amplitude de probabilité pour la particule quantique d'arriver sur la première lame séparatrice est a ; l'amplitude de probabilité de la transmission est ta, celle de la réflexion est ra. La seconde lame séparatrice recombine les faisceaux réfléchis et transmis par la première lame. On suppose que la phase du faisceau supérieur peut être modifiée de façon contrôlée, et que ce faisceau subit un changement de phase δ variable : par exemple, dans le cas d'un photon, grâce à un modulateur électro-optique (MEO). Dans ces conditions, l'amplitude de probabilité de détection par le détecteur D_1 à la sortie du second diviseur de faisceau est

$$a_1 = atr\left(1 + \mathrm{e}^{\mathrm{i}\delta}\right)$$

tandis que celle de détection par D_2 est

$$a_2 = a\left(r^2 + t^2\mathrm{e}^{\mathrm{i}\delta}\right)$$

Les probabilités p_1 et p_2 de détection par D_1 et D_2 sont donc

$$\begin{aligned} \mathsf{p}_1 &= |a_1|^2 = 2|a|^2\,|tr|^2(1 + \cos\delta) \\ \mathsf{p}_2 &= |a_2|^2 = |a|^2(|r|^4 + |t|^4 - 2|tr|^2\cos\delta) \end{aligned} \tag{1.36}$$

On vérifie que la probabilité totale de détection est bien $|a|^2$: $\mathsf{p}_1 + \mathsf{p}_2 = |a|^2$. Le phénomène d'interférences est mis en évidence en faisant varier δ : d'après (1.36), le taux de comptage du détecteur D_1, par exemple, varie sinusoïdalement en fonction de δ.

Les formules se simplifient lorsque les lames séparatrices sont équilibrées, c'est-à-dire lorsque les intensités transmises et réfléchies sont identiques : $|t| = |r| = 1/\sqrt{2}$. Dans ce cas, (1.36) devient

$$\mathsf{p}_1 = \frac{1}{2}\,|a|^2(1+\cos\delta)$$
$$\mathsf{p}_2 = \frac{1}{2}\,|a|^2(1-\cos\delta) \tag{1.37}$$

Une expérience dite "à choix retardé" a permis de montrer que la seconde lame séparatrice LS_2 peut être mise en place pendant que le photon est en vol entre LS_1 et LS_2 sans affecter la figure d'interférence : Jacques *et al.* [2007], et exercice 3.3.5. Le photon ne "choisit" pas un trajet dans l'interféromètre en franchissant LS_1.

1.5 Niveaux d'énergie

Cette section a pour objectif de définir la notion de niveau d'énergie, en rappelant d'abord la notion classique. En nous appuyant sur l'atome de Bohr, nous pourrons passer simplement à la notion quantique, puis nous examinerons les transitions radiatives entre niveaux.

1.5.1 Niveaux d'énergie en mécanique classique et modèles classiques de l'atome

Considérons une particule classique se déplaçant pour simplifier sur une droite choisie comme axe des x, et dont l'énergie potentielle est $U(x)$. En mécanique quantique, on appelle en général $U(x)$ *le potentiel.* Il est bien connu que l'énergie mécanique E, somme de l'énergie cinétique K et de l'énergie potentielle U est une constante : $E = K + U =$cste. Supposons que l'énergie potentielle a la forme de la figure 1.15, celle d'un "puits de potentiel" : elle tend vers une même constante pour $x \to \pm\infty$. Il sera commode de fixer le zéro d'énergie de telle sorte que $E = 0$ pour une particule d'énergie cinétique nulle à l'infini.

Deux situations sont possibles :

(i) La particule possède une énergie $E > 0$; alors, si elle part par exemple de $x = -\infty$, elle est d'abord accélérée puis freinée au passage du puits de potentiel et elle rejoint $x = +\infty$ avec une vitesse finale égale à sa vitesse initiale. On dit que la particule est dans un *état de diffusion.*

(ii) L'énergie est négative : $U_0 < E < 0$. Alors la particule ne peut pas sortir du puits : elle effectue des allers-retours dans le puits entre les deux points x_1 et x_2 qui vérifient $E = U(x_1) = U(x_2)$. Elle est confinée dans une région finie de l'axe des x, $x_1 \leq x \leq x_2$, et se trouve dans un

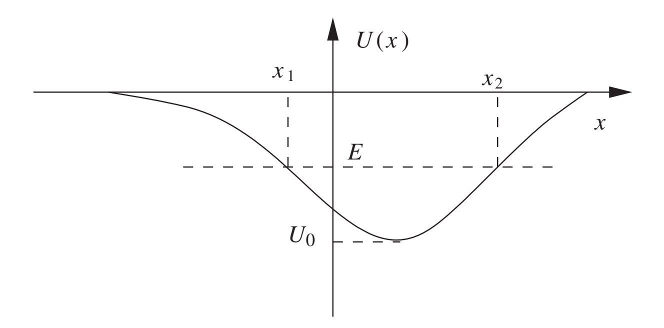


FIG. 1.15 – Puits de potentiel.

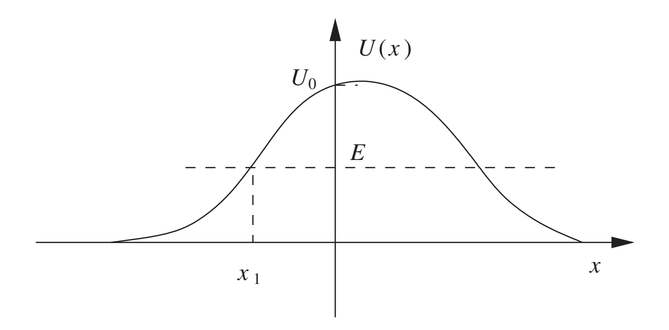


FIG. 1.16 – Barrière de potentiel.

état lié. Les points x_1 et x_2 sont appelés *points de rebroussement* de la trajectoire.

Lorsque l'énergie potentielle est positive[33] (figure 1.16), on a affaire à une “barrière de potentiel” : dans ce cas, $E > 0$ et on observe seulement des états de diffusion. Si $E < U_0$, la particule partant de $x = -\infty$ est d'abord freinée, puis elle repart en arrière au point x_1 qui vérifie $U(x_1) = E$: elle rebondit sur la barrière de potentiel. Si $E > U_0$, la particule franchit la barrière de potentiel et rejoint $x = +\infty$ en retrouvant sa vitesse initiale.

En mécanique classique, l'énergie d'un état lié peut prendre toutes les valeurs possibles entre U_0 et 0. En mécanique quantique, cette énergie ne peut prendre que des valeurs discrètes. En revanche, comme en mécanique classique, l'énergie d'un état de diffusion est arbitraire. Cependant, on trouvera (sections 8.3 et 12.4.5) des différences notables avec la mécanique classique. Par exemple, la particule peut franchir une barrière de potentiel même si

33. Naturellement on peut envisager des situations plus complexes que celles de la figure (par exemple des double-puits) ; nous nous contentons de décrire les cas les plus simples.

$E < U_0$: c'est "l'effet tunnel". Inversement, elle peut repartir en arrière même si $E > U_0$, ce que l'on appelle réflexion quantique.

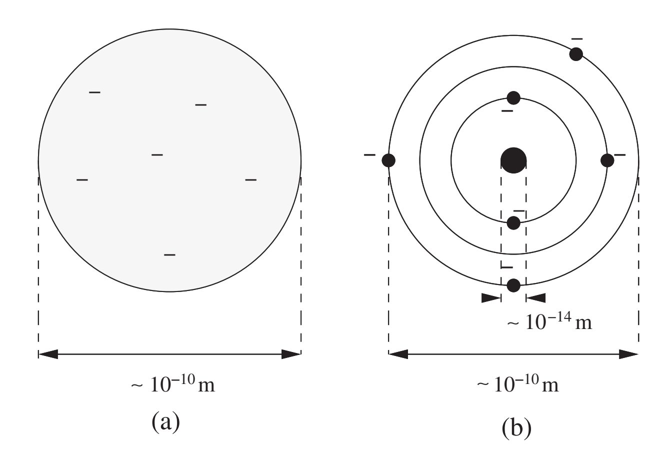


FIG. 1.17 – Modèles d'atome : (a) Thomson : les électrons sont situés dans une distribution de charge positive continue. (b) Rutherford : les électrons décrivent des orbites autour du noyau.

Appliquons ces considérations de mécanique classique aux atomes : le premier modèle d'atome fut proposé par Thomson (figure 1.17a), qui le représentait par une sphère uniformément chargée positivement, avec des électrons en mouvement dans cette distribution de charge. Un résultat élémentaire d'électromagnétisme montre que les électrons se trouvent dans un potentiel harmonique, et leur niveau d'énergie fondamental (stable) est celui où ils sont immobiles au fond du puits de potentiel ; les états excités correspondent à des vibrations autour de la position d'équilibre. Ce modèle fut éliminé[34] par les expériences de Geiger et Marsden, qui montrèrent que la diffusion de particules α (noyaux d'^{4}He) par des atomes était incompatible avec ce modèle. Rutherford déduisit de ces expériences l'existence du noyau atomique, de dimension inférieure à 10 fm, et proposa le modèle planétaire de l'atome (figure 1.17b) : les électrons tournent autour du noyau, tout comme les planètes tournent autour du Soleil, l'attraction gravitationnelle étant remplacée par l'attraction coulombienne. Ce modèle présente deux défauts majeurs, non indépendants : aucune échelle ne fixe les dimensions de l'atome et l'atome est instable, car les électrons sur orbite rayonnent et finissent par tomber sur le noyau. Dans ce processus, un spectre continu de fréquences est émis. Au contraire, les résultats expérimentaux de la fin du XIXe siècle montraient que (figure 1.18) :

- les fréquences du rayonnement émis ou absorbé par un atome sont discrètes, elles s'expriment en fonction de deux indices entiers n et m et peuvent s'écrire comme des différences : $\omega_{nm} = A_n - A_m$;

34. Pourtant, ce modèle fait encore le bonheur des physiciens atomistes...

- il existe une configuration fondamentale de l'atome où celui-ci ne rayonne pas.

Ces résultats suggéraient que l'atome émettait ou absorbait un photon en passant d'un niveau à un autre, la fréquence ω_{nm} du photon étant donné par $(E_n > E_m)$

$$\hbar\omega_{nm} = E_n - E_m \tag{1.38}$$

Les fréquences ω_{nm} sont appelées *fréquences de Bohr*. Suivant ces arguments,

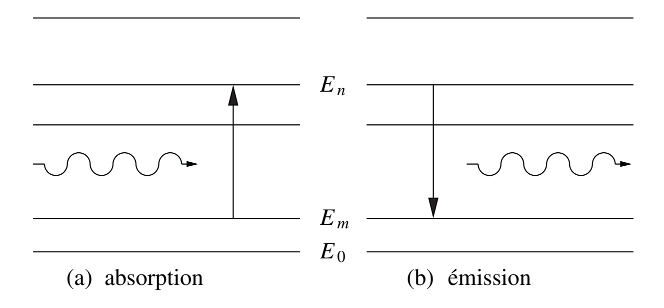


FIG. 1.18 – Émission et absorption de rayonnement entre deux niveaux E_n et E_m.

seuls certains niveaux repérés par un indice discret peuvent exister : c'est la *quantification des niveaux d'énergie*.

1.5.2 L'atome de Bohr

Pour expliquer cette quantification, Bohr plaque sur la mécanique classique et l'atome de Rutherford une règle *ad hoc* de quantification. Nous utiliserons une version légèrement différente de l'argument original de Bohr. Considérant pour simplifier l'atome d'hydrogène et une orbite circulaire de rayon a pour l'électron de masse m_e et de charge q_e, nous postulerons que le périmètre de l'orbite $2\pi a$ doit être un multiple entier de la longueur d'onde de de Broglie λ

$$2\pi a = n\lambda \qquad n = 1, 2, \ \ldots \tag{1.39}$$

Intuitivement, ceci peut s'expliquer, car cette condition revient à exiger que la phase de l'onde de de Broglie revienne à sa valeur initiale après un tour complet, et on forme ainsi une onde stationnaire. On déduit de (1.39) et (1.26)

$$2\pi a = n\,\frac{h}{p} = \frac{nh}{m_e v}$$

D'après la loi de Newton

$$\frac{m_e v^2}{a} = \frac{q_e^2}{4\pi\varepsilon_0 a^2} = \frac{e^2}{a^2} \quad \text{d'où} \quad v^2 = \frac{e^2}{m_e a}$$

où nous avons défini la quantité $e^2 = q_e^2/(4\pi\varepsilon_0)$. En éliminant la vitesse v entre ces deux équations, on obtient le rayon de l'orbite

$$a = \frac{n^2\hbar^2}{m_e e^2} \tag{1.40}$$

Le cas $n = 1$ correspond à l'orbite de plus petit rayon, et ce rayon, appelé *rayon de Bohr*, est noté a_0

$$\boxed{a_0 = \frac{\hbar^2}{m_e e^2} \simeq 0.53 \ \ \text{Å}} \tag{1.41}$$

L'énergie du niveau étiqueté par n est

$$E_n = \frac{1}{2}\, m_e v^2 - \frac{e^2}{a} = -\frac{e^2}{2a} = -\frac{m_e e^4}{2n^2\hbar^2} = -\frac{R_\infty}{n^2}$$

Les niveaux d'énergie E_n s'expriment en fonction de la *constante de Rydberg*[35] R_∞

$$\boxed{R_\infty = \frac{m_e e^4}{2\hbar^2} \simeq 13.6 \ \ \text{eV}} \tag{1.42}$$

sous la forme

$$\boxed{E_n = -\frac{R_\infty}{n^2}} \tag{1.43}$$

Cette formule donne le *spectre* (des niveaux d'énergie) de l'atome d'hydrogène. Le niveau fondamental correspond à $n = 1$ et l'énergie d'ionisation de l'atome d'hydrogène est R_∞. Les photons émis par l'atome d'hydrogène ont des fréquences

$$\hbar\omega_{nm} = -R_\infty \left(\frac{1}{n^2} - \frac{1}{m^2}\right) \quad n > m \tag{1.44}$$

en parfait accord avec les données spectroscopiques sur l'hydrogène. La simplicité avec laquelle la théorie de Bohr permet de calculer le spectre de l'atome d'hydrogène ne doit cependant pas masquer son caractère artificiel.

La généralisation par Sommerfeld de la théorie de Bohr consiste à postuler la relation

$$\int p_i \, \mathrm{d}q_i = nh \tag{1.45}$$

où q_i et p_i sont les coordonnées et les moments conjugués au sens de la mécanique classique et n un entier ≥ 1. Cependant, on sait aujourd'hui que les conditions (1.45) ne sont valables que pour certains systèmes très particuliers, les systèmes intégrables (§ 12.4.4) et pour n grand, sauf exception. La théorie de Bohr-Sommerfeld est incapable de décrire les atomes à plusieurs électrons, ainsi que les états de diffusion. Le succès de la théorie de Bohr pour l'atome d'hydrogène est un hasard heureux !

35. La raison pour l'indice ∞ est la suivante : la théorie exposée ici suppose le proton infiniment lourd. Tenir compte de la masse finie m_p du proton modifie R_∞ en $R_\infty(1/(1 + m_e/m_p))$: *cf.* exercice 1.6.5.

1.5.3 Ordres de grandeur en physique atomique

Les unités MKS, adaptées au monde à notre échelle, sont malcommodes en physique atomique. *A priori* doivent intervenir les constantes fondamentales $\hbar$ et c, ainsi que la masse m_e de l'électron ; le proton peut être considéré comme infiniment lourd, ou mieux la masse de l'électron peut être remplacée par la masse réduite (*cf.* note 35). Rappelons la valeur de ces constantes, avec une précision de $\sim 10^{-3}$ qui nous suffira pour les applications numériques

$$\begin{aligned} \hbar &= 1.055 \times 10^{-34}\,\text{J.s} \\ c &= 3 \times 10^{8}\,\text{m.s}^{-1} \\ m_e &= 0.911 \times 10^{-30}\,\text{kg} \end{aligned}$$

À partir de ces constantes, on peut fabriquer des unités naturelles

- Unité de longueur[36] : $\dfrac{\hbar}{m_e c} = 3.86 \times 10^{-13}\,\text{m}$
- Unité de temps : $\dfrac{\hbar}{m_e c^2} = 1.29 \times 10^{-21}\,\text{s}$
- Unité d'énergie : $m_e c^2 = 5.11 \times 10^{5}\,\text{eV}$

Ces unités sont déjà plus proches que les unités MKS des ordres de grandeur caractéristiques de la physique atomique, mais il manque encore quelques ordres de grandeur. En fait, on doit faire intervenir une quantité mesurant l'intensité de la force, la constante de couplage $e^2 = q_e^2/(4\pi\varepsilon_0)$. À partir de $\hbar$, c et e^2, on forme une quantité sans dimension, la *constante de structure fine*[37] α

$$\alpha = \frac{e^2}{\hbar c} = \frac{q_e^2}{4\pi\varepsilon_0 \hbar c} \simeq \frac{1}{137} \tag{1.46}$$

Les relations entre unités atomiques et unités naturelles sont maintenant faciles à obtenir ; pour le rayon de Bohr, unité de longueur naturelle de la

---

36. Appelée *longueur d'onde Compton* de l'électron.

37. Cette terminologie est utilisée pour des raisons historiques et prête à confusion ; il vaudrait bien mieux utiliser "constante atomique". α est la *constante de couplage de l'électrodynamique* ; "constante" est à mettre entre guillemets, car α n'est pas vraiment une constante, en raison de propriétés subtiles de la théorie quantique des champs. Les fluctuations quantiques du champ électron-positron ont un effet d'écran sur la charge : en raison des paires (virtuelles) électron-positron, la charge d'une particule testée à grande distance est plus petite que la même charge testée à courte distance. D'après l'inégalité de Heisenberg (1.30), courte distance implique grande impulsion, et donc grande énergie : pour explorer à courte distance, il faut utiliser des particules de haute énergie. On peut donc conclure que la constante de structure fine est une fonction croissante de l'énergie, et de fait, si on se place à des échelles d'énergie de l'ordre de l'énergie au repos du boson Z^0, $m_Z c^2 \simeq 90\,\text{GeV}$, alors $\alpha \simeq 1/129$, au lieu de la valeur de basse énergie $\alpha \simeq 1/137$. La procédure de renormalisation, qui élimine les infinis, permet de choisir une échelle d'énergie (ou de distance) arbitraire pour définir α. En résumé, α dépend de l'échelle d'énergie caractéristique du processus étudié, et aussi du détail de la procédure de renormalisation (*cf.* note 13). Cette dépendance en énergie de α est visible depuis quelques années dans les expériences de précision de la physique des hautes énergies. Voir aussi l'exercice 15.4.3.

physique atomique, on obtient

$$a_0 = \frac{\hbar^2}{m_e e^2} = \frac{\hbar c}{e^2}\,\frac{\hbar}{m_e c} = \frac{1}{\alpha}\,\frac{\hbar}{m_e c} \simeq 0.53\ \text{Å} = 0.053\ \text{nm} \tag{1.47}$$

Le Rydberg, unité naturelle d'énergie en physique atomique est relié à $m_e c^2$ par

$$R_\infty = \frac{1}{2}\,\frac{m_e e^4}{\hbar^2} = \frac{1}{2}\left(\frac{e^2}{\hbar c}\right)^2 m_e c^2 = \frac{1}{2}\,\alpha^2 m_e c^2 \simeq 13.6\ \text{eV} \tag{1.48}$$

La vitesse de l'électron sur l'orbite du niveau fondamental est $v = \alpha c = e^2/\hbar$, et la période de cette orbite, qui est l'unité de temps atomique, vaut

$$T = \frac{2\pi a_0}{v} = 2\pi\,\frac{1}{\alpha}\,\frac{\hbar}{m_e c}\,\frac{1}{\alpha c} = \frac{2\pi}{\alpha^2}\,\frac{\hbar}{m_e c^2} \simeq 1.5 \times 10^{-16}\ \text{s} \tag{1.49}$$

Les équations (1.47−1.49) montrent qu'unités naturelles et unités atomiques sont reliées par des puissances de α.

Comme dernier exemple, donnons une estimation de la vie moyenne d'un atome dans un état excité. Nous allons utiliser une image classique, en admettant que l'état excité de l'atome est modélisé par un électron sur une orbite de rayon[38] a autour du noyau, image que nous allons corriger en la complétant judicieusement le moment venu par des considérations quantiques : c'est ce que l'on appelle un *raisonnement semi-classique*. Un calcul d'électromagnétisme classique montre qu'un électron sur une orbite circulaire parcourue avec une vitesse $v = \omega a \ll c$ rayonne une puissance

$$P = \frac{2}{3c^3}\,e^2 a^2 \omega^4 = \frac{2}{3}\left(\frac{e^2}{\hbar c}\right)\frac{a^2 \hbar \omega^4}{c^2} \sim \alpha \omega^2 \hbar \left(\frac{a\omega}{c}\right)^2 \tag{1.50}$$

C'est ici qu'intervient le raisonnement quantique : l'atome émettra un photon quand il aura accumulé une énergie $\sim \hbar\omega$, ce qui va lui prendre un temps τ qui sera précisément la *vie moyenne de l'état excité*

$$\frac{1}{\tau} \sim \frac{P}{\hbar\omega} \sim \alpha\omega\left(\frac{a\omega}{c}\right)^2 \tag{1.51}$$

Mais nous avons vu que $a\omega/c = v/c \sim \alpha$, et le rapport de la période T à la vie moyenne τ est

$$\frac{T}{\tau} \sim \frac{1}{\tau\omega} \sim \alpha^3 \sim 10^{-6} \tag{1.52}$$

L'orbite est parcourue environ un million de fois avant l'émission d'un photon : un état excité est donc bien défini. Pour l'état fondamental de l'atome

38. On peut aussi assimiler l'atome à un dipôle oscillant de fréquence ω, comme dans le modèle de Thomson. La seule différence est que le facteur 2/3 dans (1.50) devient 1/3, ce qui est sans influence sur les ordres de grandeur.

d'hydrogène dont l'énergie est $\sim 10\,\text{eV}$, nous avons vu que $T \sim 10^{-16}\,\text{s}$; pour l'électron externe d'un alcalin avec une énergie $\sim 1\,\text{eV}$, nous avons plutôt $T \sim 10^{-15}\,\text{s}$, et l'ordre de grandeur typique de la vie moyenne d'un état excité est $\sim 10^{-7}-10^{-9}\,\text{s}$. Par exemple, le premier niveau excité du rubidium a une vie moyenne de $2.7 \times 10^{-8}\,\text{s}$.

Les raisonnements utilisés dans cette section ont le mérite de la simplicité, mais ils ne sont pas satisfaisants. Ils consistent à plaquer arbitrairement une contrainte quantique sur un raisonnement classique, au moment où celui-ci devient intenable, et le lecteur pourra estimer à juste titre qu'il n'est pas convaincu par ce type de raisonnement. Il est donc indispensable de passer à une théorie entièrement nouvelle, qui ne soit plus guidée par la physique classique, mais qui *développe son propre cadre de façon autonome, sans référence à la physique classique.*

1.6 Exercices

1.6.1 Ordres de grandeur

1. On se propose d'explorer des distances de l'ordre de la taille d'un atome, soit $\sim 1\,\text{Å}$ (0.1 nm), avec des photons, des neutrons ou des électrons. Quel sera en eV l'ordre de grandeur de l'énergie de ces particules ?

2. Lorsque la longueur d'onde λ d'une onde sonore est grande par rapport au pas du réseau cristallin où se propage la vibration, la fréquence ω de cette onde sonore est linéaire dans le vecteur d'onde $k = 2\pi/\lambda$: $\omega = c_s k$, où c_s est la vitesse du son. Dans le cas de l'acier, $c_s \simeq 5 \times 10^3\,\text{m.s}^{-1}$. Quelle est l'énergie $\hbar\omega$ d'une onde sonore pour $k = 1\,\text{nm}^{-1}$? La particule analogue du photon pour les ondes sonores est appelée *phonon*, et $\hbar\omega$ est l'énergie d'un phonon. Sachant qu'un phonon peut être créé par collision inélastique sur le cristal, utiliseriez-vous des neutrons ou des photons pour étudier les phonons ?

3. Dans une expérience d'interférences avec des fullerènes C_{60}, qui sont aujourd'hui les plus gros objets avec lesquels on a vérifé le comportement ondulatoire[39], la vitesse moyenne des molécules est de 220 m.s^{-1}. Quelle est leur longueur d'onde de de Broglie ? Comment se compare-t-elle aux dimensions de la molécule ?

4. Une molécule diatomique est formée de deux atomes de masse M_1 et M_2 ; elle a la forme d'une haltère. Les deux noyaux atomiques sont distants de $r_0 = ba_0$, où a_0 est le rayon de Bohr (1.41) et b un coefficient numérique ~ 1. On suppose que la molécule tourne autour de son centre d'inertie suivant un axe perpendiculaire à la droite joignant les noyaux, appelée droite des noyaux. Montrer que son moment d'inertie est $I = \mu r_0^2$ où $\mu = M_1M_2/(M_1+M_2)$ est la masse réduite. On suppose que son moment angulaire est $\hbar$. Quelle est alors la vitesse angulaire de rotation et quelle est l'énergie ε_{rot} correspondante ?

39. Arndt *et al.* [1999] et [2005].

Montrer que cette énergie est proportionnelle à $(m_e/\mu)R_\infty$ où m_e est la masse de l'électron et $R_\infty = m_e e^4/(2\hbar^2) = e^2/(2a_0)$.

5. La molécule peut aussi vibrer le long de la droite des noyaux autour de sa position d'équilibre $r = r_0$, la force de rappel étant de la forme $-K(r-r_0)$, avec $Kr_0^2 = cR_\infty$ où c est un coefficient numérique ~ 1. Quelles sont la fréquence de vibration ω_v et l'énergie $\hbar\omega_v$ correspondante ? Montrer que cette énergie est proportionnelle à $\sqrt{m_e/\mu}\, R_\infty$. Exemple : la molécule de HCl^{35}, où les valeurs expérimentales sont $r_0 = 1.27$ Å (0.127 nm), $\varepsilon_{\rm rot} = 1.3 \times 10^{-3}$ eV, $\hbar\omega_v = 0.36$ eV. Calculer les valeurs numériques de b et c. Quelle serait la longueur d'onde d'un photon ayant l'énergie $\varepsilon_{\rm rot}$, $\hbar\omega_v$? Dans quels domaines se trouvent ces longueurs d'onde ?

6. L'absence d'une théorie quantique de la gravitation oblige à limiter toute théorie à des énergies plus petites que E_P, l'énergie de Planck. Par un argument dimensionnel, construire E_P en fonction de la constante de gravitation G (1.5), $\hbar$ et c et donner sa valeur numérique. Quelle est la longueur correspondante, ou longueur de Planck l_P ?

1.6.2 Le corps noir

1. Démontrer l'équation (note 22)

$$\int \mathrm{d}x\mathrm{d}p\, \delta\left(E - \frac{p^2}{2m} - \frac{1}{2}\, m\omega^2 x^2\right) f(E) = \frac{2\pi}{\omega}\, f(E)$$

2. On se propose de relier la densité d'énergie par unité de fréquence $\epsilon(\omega, T)$ à la puissance émise $u(\omega, T)$: équation (1.15). On considère une enceinte portée à la température T (figure 1.4). Soit $\tilde{\epsilon}(k, T)\mathrm{d}^3k$ la densité d'énergie dans d^3k autour de $\vec{k}$, qui ne dépend que de $k = |\vec{k}|$. Montrer que

$$\tilde{\epsilon}(k, T) = \frac{c}{4\pi k^2}\, \epsilon(\omega, T)$$

Le vecteur de Poynting pour une onde s'échappant de l'enceinte avec un vecteur d'onde $\vec{k}$ est $c\tilde{\epsilon}(k, T)\hat{k}$. Montrer que le flux du vecteur de Poynting à travers une ouverture d'aire $\mathcal{S}$ est

$$\Phi = \frac{1}{4}\, c\mathcal{S} \int_0^\infty \epsilon(\omega, T)\mathrm{d}\omega$$

et en déduire (1.15).

3. Montrer par analyse dimensionnelle qu'en physique classique on doit avoir pour la densité d'énergie du corps noir

$$\epsilon(T) = A(k_{\rm B}T)c^{-3} \int_0^\infty \omega^2\, \mathrm{d}\omega$$

où A est un coefficient numérique.

4. Chaque mode $\vec{k}$ du champ électromagnétique dans l'enceinte est un oscillateur harmonique. En mécanique statistique classique, l'énergie d'un tel mode est $2k_\mathrm{B}T$ (pourquoi ce facteur 2 ?). Montrer que la densité d'énergie dans l'enceinte est

$$\epsilon(T) = \frac{1}{\pi^2}\,(k_\mathrm{B}T)c^{-3}\int_0^\infty \omega^2\,\mathrm{d}\omega$$

et en déduire A.

5. Démontrer (1.22) et vérifier que l'on retrouve l'expression classique pour $\hbar\omega \ll k_\mathrm{B}T$, c'est-à-dire pour une température suffisamment grande à ω fixé. Ceci est un résultat très général : *l'approximation classique est valable à haute température.*

1.6.3 Inégalités de Heisenberg

Dans l'expérience théorique de la figure 1.12, montrer que l'impulsion δp_x communiquée à l'écran vaut $pa/(2D)$, où a est la distance entre les fentes F_1 et F_2 (figure 1.12) et p l'impulsion des photons. La détermination de la trajectoire implique que $\Delta p_x \ll \delta p_x$, où Δp_x est la dispersion sur l'impulsion *initiale* de l'écran. Quelle est alors la dispersion Δx sur sur la position de F_0 ? En déduire que les interférences sont alors détruites. Voir Wootters et Zurek [1979].

1.6.4 Diffraction de neutrons par un cristal

La diffraction des neutrons est une des principales techniques d'analyse de la structure des cristaux. On considère pour simplifier un cristal bidimensionnel composé d'atomes identiques, les vecteurs d'onde étant situés dans le plan du cristal[40]. Les atomes du cristal se trouvent aux points du réseau (figure 1.19)

$$\vec{r}_i = na\hat{x} + mb\hat{y} \qquad n = 0, 1, \ \ldots N-1; \qquad m = 0, 1, \ \ldots M-1$$

Les neutrons interagissent avec les noyaux des atomes[41] par une interaction de type nucléaire. On appelle $f(\theta)$ l'amplitude de probabilité pour qu'un neutron d'impulsion $\hbar\vec{k}$ soit diffusé dans la direction $\hat{k}'$ par un atome situé à l'origine des coordonnées, θ étant l'angle entre $\hat{k}$ et $\hat{k}'$. Comme l'énergie des neutrons est très faible, $\sim 0.01\,\mathrm{eV}$, $f(\theta)$ est indépendant de θ (§ 13.2.4) : $f(\theta) = f$. La collision entre le neutron et le noyau atomique est élastique et l'état du cristal est inchangé dans la collision : il est impossible de savoir quel atome a diffusé un neutron.

---

40. On peut aussi envisager une diffusion $3D$ par un cristal $2D$: *cf.* Wichman [1974], chapitre 5, ce qui donne un modèle pour la diffraction par la surface d'un cristal.

41. Il existe aussi une interaction entre le moment magnétique du neutron et le magnétisme de l'atome, qui joue un rôle très important pour l'étude du magnétisme, mais qui ne nous concerne pas dans ce problème.

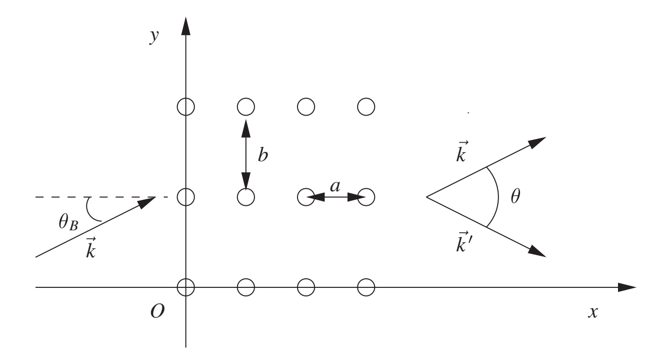


FIG. 1.19 – Diffraction de neutrons par un cristal.

1. Montrer que l'amplitude de diffusion par un atome situé au point $\vec{r}_i$ est

$$f_i = f\,\mathrm{e}^{\mathrm{i}(\vec{k}-\vec{k}')\cdot\vec{r}_i} = f\,\mathrm{e}^{-\mathrm{i}\vec{q}\cdot\vec{r}_i}$$

avec $\vec{q} = \vec{k}' - \vec{k}$.

2. Montrer que l'amplitude de diffusion $f_{\rm tot}$ par le cristal est de la forme

$$f_{\rm tot} = fF(aq_x, bq_y)$$

la fonction $F(aq_x, bq_y)$ étant donnée par

$$F(aq_x, bq_y) = \exp\left(-\mathrm{i}\frac{aq_x(N-1)}{2}\right)\exp\left(-\mathrm{i}\frac{bq_y(M-1)}{2}\right) \times\left[\frac{\sin(aq_xN/2)}{\sin(aq_x/2)}\right]\left[\frac{\sin(bq_yM/2)}{\sin(bq_y/2)}\right]$$

3. Montrer que pour $N, M \gg 1$, la probabilité de diffusion est proportionnelle à $(NM)^2$ lorsque $\vec{q}$ a pour composantes

$$q_x = \frac{2\pi n_x}{a} \qquad q_y = \frac{2\pi n_y}{b}$$

les nombres n_x et n_y étant des nombres entiers : lorsque les composantes de $\vec{q}$ sont de cette forme, on dit que $\vec{q}$ appartient au *réseau réciproque* du réseau cristallin. On obtient des maxima de diffraction si $\vec{q}$ est un vecteur du réseau réciproque. Quelle est la largeur du pic de diffraction autour d'un maximum ? En déduire que l'intensité dans le pic est proportionnelle à NM.

4. On doit tenir compte du caractère élastique de la diffusion. Montrer que la condition de diffusion élastique est

$$2\vec{k}\cdot\vec{q} + q^2 = 0$$

Un vecteur du réseau réciproque ne donnera un maximum de diffraction que si cette condition est vérifiée. Pour une longueur d'onde fixée, cette condition ne

pourra être satisfaite que si l'angle d'incidence prend des valeurs particulières, appelées *angles de Bragg* θ_B. Une étude simple est possible si $n_x = 0$. Montrer que, dans ce cas, un angle d'incidence θ_B donne lieu à diffraction pour

$$\sin\theta_B = \frac{\pi n}{bk}, \qquad n = 1, 2, \cdots$$

Dans le cas général, il est commode de donner une interprétation géométrique de la condition de Bragg : l'extrémité du vecteur $\vec{k}$ étant située sur un point du réseau réciproque, on trace un cercle de rayon k. Si ce cercle passe par un autre point du réseau réciproque, alors on obtiendra un maximum de diffraction. En général, un faisceau de neutrons incident sur un cristal ne donnera pas de pic de diffraction. Il faut choisir convenablement l'angle d'incidence et/ou la longueur d'onde. Pourquoi ce phénomène ne se produit-il pas pour la diffraction par un réseau à une dimension ? Que se passerait-il s'il y avait seulement la première rangée verticale d'atomes sur la droite $y = 0$?

5. On suppose maintenant que le cristal est formé de deux types d'atomes. Le motif élémentaire, ou *maille* du cristal, est formé de la façon suivante : deux atomes de type 1 sont situés respectivement en

$$\vec{r}_1 = 0 \quad \text{et} \quad \vec{r}_1{}' = a\hat{x} + b\hat{y}$$

et deux atomes de type 2 en

$$\vec{r}_2 = a\hat{x} \quad \text{et} \quad \vec{r}_2{}' = b\hat{y}$$

La maille se répète avec une périodicité $2a$ dans la direction x et $2b$ dans la direction y. Soit f_1 (f_2) l'amplitude de diffusion d'un neutron par un atome de type 1 (2) situé à l'origine des coordonnées ; on pourra prendre f_1 et f_2 réels. Si NM est le nombre de mailles, montrer que l'amplitude de diffusion par le cristal est proportionnelle à $F(2aq_x, 2bq_y)$. Déterminer le facteur de proportionnalité en fonction de f_1 et f_2. Montrer que si q_x et q_y correspondent à un maximum de diffraction, ce facteur de proportionnalité vaut

$$f_1\left[1 + (-1)^{n_x+n_y}\right] + f_2\left[(-1)^{n_x} + (-1)^{n_y}\right]$$

Discuter le résultat en fonction de la parité de n_x et de n_y.

6. Les atomes 1 et 2 forment un alliage[42] : à basse température les atomes sont dans la configuration de la question **5**, mais au-dessus d'une certaine température chaque atome a une probabilité de 50 % d'occuper un site quelconque et tous les sites sont équivalents. Comment va évoluer la figure de diffraction ?

42. Un exemple du phénomène décrit dans cette question est donné par le bronze, pour une proportion de 50 % de cuivre et de 50 % de zinc.

1.6.5 Atomes hydrogénoïdes

Calculer en fonction de R_∞ le niveau d'énergie fondamental de l'atome d'hydrogène ordinaire, de l'atome de deutérium et de l'atome d'hélium une fois ionisé en tenant compte de ce que la masse des nucléons est finie. Suggestion : quelles sont les masses réduites ?

1.6.6 Interféromètre à neutrons et gravité

Un interféromètre à neutrons (figure 1.20), taillé dans un monocristal de silicium de quelques cm de côté, est fondé sur le même principe que celui de Mach-Zehnder (§ 1.4.5) : le faisceau incident supposé monochromatique (c'est-à-dire de longueur d'onde fixée) arrive sur une première lame séparatrice en A, l'angle d'incidence et la longueur d'onde étant choisies de telle sorte que l'on obtienne un pic de diffraction (voir exercice 1.6.4, question **4**) : l'angle d'incidence est un angle de Bragg θ_B. Une partie du faisceau est transmise dans le faisceau I avec une amplitude de probabilité t et l'autre partie est réfractée dans le faisceau II avec une amplitude de probabilité r. Ces amplitudes vérifient $|t|^2 + |r|^2 = 1$. Les faisceaux I et II arrivent sur une seconde lame séparatrice, respectivement en B, et en D, et les parties réfractées de I et II sont recombinées par une troisième lame séparatrice en C. Les neutrons sont détectés par deux compteurs D_1 et D_2. Sur le trajet II, les neutrons subissent un déphasage δ qui peut avoir diverses origines (différence de longueur entre les trajets, gravité, passage dans un champ magnétique, etc.), et l'objectif de l'interférométrie neutronique est précisément de mesurer ce déphasage.

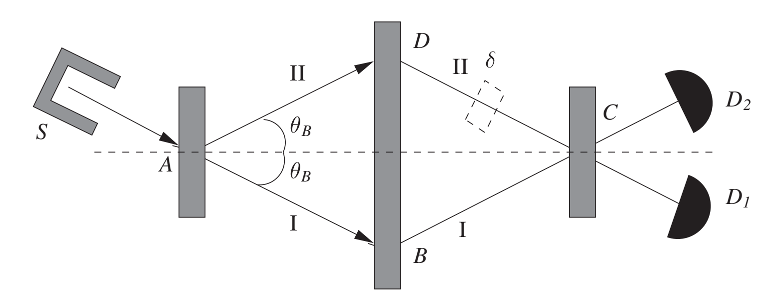


Fig. 1.20 – Interféromètre à neutrons. S est la source de neutrons.

1. Montrer que l'amplitude de probabilité a_1 pour qu'un neutron arrive sur D_1 est

$$a_1 = a_0(\mathrm{e}^{\mathrm{i}\delta}\, trr + rrt)$$

et que la probabilité de détection par D_1 est

$$\mathsf{p}_1 = 2|a_0|^2|t|^2|r|^4(1+\cos\delta) = A(1+\cos\delta)$$

a_0 étant l'amplitude incidente sur le premier cristal.

2. Quelle sont l'amplitude a_2 d'arrivée d'un neutron sur le détecteur D_2 en fonction de r, t et a_0 et la probabilité p_2 correspondante ? Pourquoi doit-on avoir $\mathsf{p}_1 + \mathsf{p}_2 = \text{cste}$? En déduire

$$\mathsf{p}_2 = B - A\cos\delta$$

Quelle est l'expression de B en fonction de t, r et a_0 ? Posant

$$t = |t|\mathrm{e}^{\,\mathrm{i}\alpha} \qquad r = |r|\mathrm{e}^{\,\mathrm{i}\beta}$$

montrer que

$$\alpha - \beta = \frac{\pi}{2} \pm n\pi \qquad n = 0, 1, 2, \ldots$$

3. On tient compte de la gravité : comment varie en fonction de l'altitude z le vecteur d'onde $k = 2\pi/\lambda$ d'un neutron lorsqu'il est placé dans un champ de pesanteur, l'accélération de la pesanteur étant g ? Comparer les valeurs numériques de l'énergie cinétique du neutron et de son énergie gravitationnelle[43] $m_n g z$ (m_n est la masse du neutron) et en déduire une approximation pour k. Le plan $ABDC$ étant initialement horizontal, on fait tourner autour de AB ce plan qui devient vertical. Montrer que cette rotation induit une différence de phase entre les deux trajets

$$\Delta\phi = \frac{m_n^2 g \mathcal{S}}{\hbar^2 k} = \frac{2\pi m_n^2 g \mathcal{S}\lambda}{h^2}$$

où $\mathcal{S}$ est l'aire du losange $ABDC$.

4. Si le plan $ABDC$ fait un angle variable θ avec la verticale, discuter qualitativement la variation de la probabilité de détection des neutrons en fonction de θ. Données numériques (Colella *et al.* [1975]) : $\lambda = 1.44\,\text{Å}$, $\mathcal{S}$=10.1 cm^2.

1.6.7 Diffusion cohérente et diffusion incohérente de neutrons par un cristal

On se propose d'étudier la diffusion de neutrons par un cristal formé de deux types de noyaux. Un site donné du cristal est occupé par un noyau de type 1 avec une probabilité p_1 ou par un noyau de type 2 avec une probabilité $p_2 = 1 - p_1$. Le nombre total de noyaux est $\mathcal{N}$, et il y a donc $p_1\mathcal{N}$ noyaux de type 1 et $p_2\mathcal{N}$ noyaux de type 2 dans le cristal. Au site i, $i = 1, \cdots, \mathcal{N}$, on associe un nombre α_i qui prend la valeur 1 si le site est occupé par un noyau de type 1 et 0 s'il est occupé par un noyau de type 2. L'ensemble $\{\alpha_i\}$ des α_i, avec $\sum_i \alpha_i = p_1\mathcal{N}$, définit une configuration du cristal. L'amplitude

43. L'énergie étant définie à une constante additive près, on fixe par convention le zéro d'énergie de la façon suivante : un neutron de vitesse nulle et d'altitude $z = 0$ a une énergie nulle.

de diffusion d'un neutron par le cristal dans la configuration $\{\alpha_i\}$ est (*cf.* l'exercice 1.6.4)

$$f_{\rm tot} = \sum_{i=1}^{\mathcal{N}} (\alpha_i f_1 + (1-\alpha_i) f_2)\, {\rm e}^{{\rm i}\vec{q}\cdot\vec{r}_i}$$

où f_1 (f_2) est l'amplitude de diffusion d'un neutron par un noyau de type 1 (2).

1. On note $\langle \bullet \rangle$ la moyenne sur toutes les configurations possibles du cristal, en supposant que les occupations des sites sont ne sont pas corrélées (par exemple l'occupation d'un site par un noyau de type 1 ne doit pas augmenter la probabilité qu'un site plus proche voisin soit aussi occupé par un noyau de type 1). Démontrer les identités

$$\langle \alpha_i \alpha_j \rangle = p_1^2 + p_1 p_2 \delta_{ij} \qquad \langle \alpha_i (1-\alpha_j) \rangle = p_1 p_2 (1-\delta_{ij})$$

2. Déduire de ces identités la moyenne sur les configurations de $|f_{\rm tot}|^2$

$$\langle |f_{\rm tot}|^2 \rangle = (p_1 f_1 + p_2 f_2)^2 \sum_{i,j} {\rm e}^{{\rm i}\vec{q}\cdot(\vec{r}_i - \vec{r}_j)} + \mathcal{N} p_1 p_2 (f_1 - f_2)^2$$

Le premier terme décrit la *diffusion cohérente* et donne lieu à des pics de diffraction. Le deuxième est proportionnel au nombre de sites et est indépendant des angles : ce terme correspond à la *diffusion incohérente.*

1.7 Bibliographie

On trouve une introduction élémentaire à la physique quantique dans Scarani [2003], Le Bellac [2010] ou dans Hey et Walters [2003]. Il est également recommandé de lire les chapitres introductifs 1 à 3 de Feynman *et al.* [1965], volume III, 1 à 5 de Wichman [1974] ainsi que les chapitres 1 à 3 de Lévy-Leblond et Balibar [1984]. Pour une introduction pédagogique et actualisée à la physique des particules élémentaires, voir Perkins [2000] ou Tully [2011] ; voir aussi l'article grand public Jacob [2002]. On trouvera une étude détaillée du rayonnement du corps noir, par exemple, dans Diu *et al.* [1990], chapitre 4 ou Le Bellac *et al.* [2004], chapitre 5. Les expériences d'interférences et de diffraction de neutrons froids ont été réalisées par Zeilinger *et al.* [1988], et les expériences d'interférences avec des atomes froids par Shimizu *et al.* [1992]. Pour la diffraction des neutrons par un cristal, on pourra se reporter à Kittel [1970], chapitre 2. Un exposé détaillé sur l'interférométrie neutronique se trouve dans le livre de Rauch et Werner [2000].

Chapitre 2

Mathématiques de la mécanique quantique I : dimension finie

Le principe de superposition est un principe fondateur de la mécanique quantique, et nous nous sommes appuyés sur ce principe pour rendre compte des interférences, par exemple dans l'expérience des fentes d'Young. La mécanique quantique est une *théorie linéaire*, et il est naturel que les espaces vectoriels y jouent un rôle fondamental. Nous verrons qu'un état physique est représenté mathématiquement par un vecteur dans un espace dont nous allons préciser les caractéristiques, et qui sera appelé *espace des états*. Un second principe fondateur, également déduit des expériences d'interférences, est l'existence d'amplitudes de probabilité. Ces amplitudes de probabilité seront représentées mathématiquement par des produits scalaires définis sur l'espace des états. En physique des ondes, l'utilisation des nombres complexes est uniquement une commodité, mais en mécanique quantique les amplitudes de probabilité sont fondamentalement des nombres complexes : le produit scalaire sera *a priori* un nombre complexe. Les propriétés physiques : impulsion, position, énergie... seront représentées par des opérateurs agissant dans l'espace des états. Dans ce chapitre, nous introduisons les propriétés essentielles des *espaces de Hilbert*, c'est-à-dire les espaces vectoriels munis d'un produit scalaire défini positif, en nous limitant au cas de la dimension finie. Cette restriction devra être levée ultérieurement, car l'espace des états est en général de dimension infinie. La théorie mathématique des espaces de Hilbert de dimension infinie est beaucoup plus complexe que celle des espaces de dimension finie, et nous renvoyons leur étude au chapitre 6. Le lecteur familier des espaces vectoriels de dimension finie et des opérateurs dans ces espaces peut passer directement au chapitre 3 après un survol des notations.

2.1 Espaces de Hilbert de dimension finie

Soit $\mathcal{H}$ un espace vectoriel de dimension N sur le corps des complexes. Nous noterons $|\varphi\rangle$, $|\chi\rangle, \ldots$ les éléments de $\mathcal{H}$. Si λ, $\mu \ldots$ sont des nombres complexes, et si $|\varphi\rangle$ et $|\chi\rangle \in \mathcal{H}$, la linéarité implique que $\lambda|\varphi\rangle \equiv |\lambda\varphi\rangle \in \mathcal{H}$ et que $(|\varphi\rangle + \lambda|\chi\rangle) \in \mathcal{H}$.

L'espace $\mathcal{H}$ est muni d'un produit scalaire défini positif, ce qui en fait un *espace de Hilbert*. Le produit scalaire[1] de deux vecteurs $|\varphi\rangle$ et $|\chi\rangle$ sera noté $\langle\chi|\varphi\rangle$; il est linéaire par rapport à $|\varphi\rangle$

$$\langle\chi|(\varphi_1 + \lambda\varphi_2)\rangle = \langle\chi|\varphi_1\rangle + \lambda\langle\chi|\varphi_2\rangle \tag{2.1}$$

et vérifie la propriété de conjugaison complexe

$$\langle\chi|\varphi\rangle = \langle\varphi|\chi\rangle^* \tag{2.2}$$

ce qui implique que $\langle\varphi|\varphi\rangle$ est un nombre réel. De (2.1) et (2.2), on déduit que le produit scalaire $\langle\chi|\varphi\rangle$ est antilinéaire par rapport à $|\chi\rangle$

$$\langle(\chi_1 + \lambda\chi_2)|\varphi\rangle = \langle\chi_1|\varphi\rangle + \lambda^*\langle\chi_2|\varphi\rangle \tag{2.3}$$

Enfin le produit scalaire est défini positif

$$\langle\varphi|\varphi\rangle = 0 \Longleftrightarrow |\varphi\rangle = 0 \tag{2.4}$$

Il sera commode de choisir dans $\mathcal{H}$ une base orthonormée de N vecteurs $\{|1\rangle, |2\rangle, \ldots, |n\rangle, \ldots, |N\rangle\}$

$$\langle n|m\rangle = \delta_{nm} \tag{2.5}$$

Tout vecteur $|\varphi\rangle$ peut se décomposer sur cette base avec des coefficients c_n qui sont les composantes de $|\varphi\rangle$ dans cette base

$$|\varphi\rangle = \sum_{n=1}^{N} c_n|n\rangle \tag{2.6}$$

Prenant le produit scalaire de (2.6) avec le vecteur de base $|m\rangle$, on trouve pour c_m

$$c_m = \langle m|\varphi\rangle \tag{2.7}$$

Si un vecteur $|\chi\rangle$ se décompose sur cette même base suivant $|\chi\rangle = \sum d_n|n\rangle$, alors le produit scalaire $\langle\chi|\varphi\rangle$ s'écrit, en utilisant (2.5)

$$\langle\chi|\varphi\rangle = \sum_{n,m=1}^{N} d_m^*\, c_n\langle m|n\rangle = \sum_{n=1}^{N} d_n^*\, c_n \tag{2.8}$$

1. Il pourra nous arriver d'utiliser la notation des mathématiciens $(\chi, \varphi) \equiv \langle\chi|\varphi\rangle$ pour le produit scalaire. Toutefois, il faut prendre garde au fait que pour les mathématiciens le produit scalaire (χ, φ) est linéaire par rapport à χ !

La *norme* de $|\varphi\rangle$, notée $||\varphi||$, est définie à partir du produit scalaire

$$||\varphi||^2 = \langle\varphi|\varphi\rangle = \sum_{n=1}^{N} |c_n|^2 \geq 0 \tag{2.9}$$

Une propriété importante du produit scalaire est l'inégalité de Schwarz

$$\boxed{|\langle\chi|\varphi\rangle|^2 \leq \quad \langle\chi|\chi\rangle\langle\varphi|\varphi\rangle = ||\chi||^2\, ||\varphi||^2} \tag{2.10}$$

L'égalité est vraie si et seulement si $|\varphi\rangle$ et $|\chi\rangle$ sont proportionnels : $|\chi\rangle = \lambda|\varphi\rangle$.
Démonstration[2]. Le théorème est vérifié si $\langle\chi|\varphi\rangle = 0$, nous pouvons donc supposer que $\langle\chi|\varphi\rangle \neq 0 \Rightarrow |\varphi\rangle \neq 0$ et $|\chi\rangle \neq 0$. D'après la propriété (2.9) de la norme

$$\langle(\varphi - \lambda\chi)|(\varphi - \lambda\chi)\rangle = ||\varphi||^2 - \lambda^*\langle\chi|\varphi\rangle - \lambda\langle\varphi|\chi\rangle + |\lambda|^2||\chi||^2 \geq 0$$

Choisissant

$$\lambda = \frac{||\varphi||^2}{\langle\varphi|\chi\rangle} \qquad \lambda^* = \frac{||\varphi||^2}{\langle\chi|\varphi\rangle}$$

on obtient

$$||\varphi||^2 - 2||\varphi||^2 + \frac{||\varphi||^4||\chi||^2}{|\langle\chi|\varphi\rangle|^2} \geq 0$$

d'où (2.10) suit immédiatement. L'égalité ne peut avoir lieu d'après (2.4) que si $|\varphi\rangle = \lambda|\chi\rangle$ et réciproquement.

2.2 Opérateurs linéaires sur $\mathcal{H}$

2.2.1 Opérateurs linéaires, hermitiens, unitaires

Un opérateur linéaire A fait correspondre au vecteur $|\varphi\rangle$ un vecteur $|A\varphi\rangle$ vérifiant la propriété de linéarité

$$|A(\varphi + \lambda\chi)\rangle = |A\varphi\rangle + \lambda|A\chi\rangle \tag{2.11}$$

Dans une base déterminée, cet opérateur est représenté par une matrice[3] d'éléments A_{mn}. En effet grâce à la linéarité et en utilisant la décomposition (2.6)

$$|A\varphi\rangle = \sum_{n=1}^{N} c_n|An\rangle$$

2. Cette démonstration se transpose immédiatement au cas où l'espace est de dimension infinie.

3. On notera que, par abus de langage, les physiciens confondent souvent l'opérateur et sa matrice représentative dans une base donnée.

on obtient les composantes d_m de $|A\varphi\rangle = \sum_m d_m|m\rangle$

$$d_m = \langle m|A\varphi\rangle = \sum_{n=1}^{N} c_n\langle m|An\rangle = \sum_{n=1}^{N} A_{mn}c_n \tag{2.12}$$

L'élément de matrice A_{mn} est donc

$$A_{mn} = \langle m|An\rangle \tag{2.13}$$

L'opérateur *conjugué hermitien* (ou adjoint) $A^\dagger$ de A est défini par

$$\langle\chi|A^\dagger\varphi\rangle = \langle A\chi|\varphi\rangle = \langle\varphi|A\chi\rangle^* \tag{2.14}$$

pour tout couple de vecteurs $|\varphi\rangle, |\chi\rangle$. On montre facilement que $A^\dagger$ est bien un opérateur linéaire. Ses éléments de matrice dans la base $\{|1\rangle, |2\rangle, \ldots, |n\rangle, \ldots, |N\rangle\}$ sont obtenus en prenant pour $|\varphi\rangle$ et $|\chi\rangle$ les vecteurs de base et $(A^\dagger)_{mn}$ vérifie

$$(A^\dagger)_{mn} = A^*_{nm} \tag{2.15}$$

Le conjugué hermitien du produit AB de deux opérateurs est $B^\dagger A^\dagger$; en effet

$$\langle\chi|(AB)^\dagger\varphi\rangle = \langle AB\chi|\varphi\rangle = \langle B\chi|A^\dagger\varphi\rangle = \langle\chi|B^\dagger A^\dagger\varphi\rangle$$

Un opérateur vérifiant $A = A^\dagger$ est appelé *hermitien*, ou *auto-adjoint*. Les deux termes sont équivalents pour les espaces de dimension finie, mais non dans le cas de la dimension infinie.

Un opérateur tel que $UU^\dagger = U^\dagger U = I$, ou de façon équivalente $U^{-1} = U^\dagger$, est appelé opérateur *unitaire* : dans toute la suite du livre, I désignera *l'opérateur identité* de l'espace de Hilbert. Dans un espace de dimension finie, une condition nécessaire et suffisante pour qu'un opérateur U soit unitaire est qu'il conserve la norme

$$||U\varphi||^2 = ||\varphi||^2 \quad \text{ou} \quad \langle U\varphi|U\varphi\rangle = \langle\varphi|\varphi\rangle \quad \forall\varphi \in \mathcal{H} \tag{2.16}$$

Démonstration. Calculons la norme carrée de $|U(\varphi+\lambda\chi)\rangle$, qui, par hypothèse, est égale à la norme carrée de $|\varphi + \lambda\chi\rangle$

$$\langle\varphi + \lambda\chi|\varphi + \lambda\chi\rangle = \langle\varphi|\varphi\rangle + |\lambda|^2\langle\chi|\chi\rangle + 2\text{Re}\,(\lambda\langle\varphi|\chi\rangle)$$

tandis que

$$\langle U(\varphi + \lambda\chi)|U(\varphi + \lambda\chi)\rangle = \langle U\varphi|U\varphi\rangle + |\lambda|^2\langle U\chi|U\chi\rangle + 2\text{Re}\,(\lambda\langle U\varphi|U\chi\rangle)$$

En retranchant la seconde des équations ci-dessus de la première

$$\text{Re}\,(\lambda\langle\varphi|\chi\rangle) = \text{Re}\,(\lambda\langle U\varphi|U\chi\rangle)$$

et en choisissant $\lambda = 1$ puis $\lambda = \text{i}$ on déduit

$$\langle U\varphi|U\chi\rangle = \langle\varphi|\chi\rangle \Rightarrow U^\dagger U = I$$

Dans un espace vectoriel de dimension finie, l'existence d'un inverse à gauche entraîne celle d'un inverse à droite, et on a également $UU^\dagger = I$. Un opérateur qui conserve la norme est une *isométrie*. Dans un espace de dimension finie, une isométrie est un opérateur unitaire.

Les opérateurs unitaires effectuent les changements de base orthonormée dans $\mathcal{H}$. Soit $|n'\rangle = |Un\rangle$, alors

$$\langle m'|n'\rangle = \langle Um|Un\rangle = \langle m|n\rangle = \delta_{mn} = \delta_{m'n'}$$

et l'ensemble des vecteurs $\{|n'\rangle\}$ forme une base orthonormée. Il faut prendre garde au fait que les composantes c_n d'un vecteur se transforment à l'aide de $U^\dagger$ (ou U^{-1})

$$c'_n = \langle n'|\varphi\rangle = \langle Un|\varphi\rangle = \langle n|U^\dagger\varphi\rangle = \sum_{m=1}^{N} U^\dagger_{nm} c_m \tag{2.17}$$

Notons également la loi de transformation des éléments de matrice

$$A'_{mn} = \langle m'|An'\rangle = \langle Um|AUn\rangle = \langle m|U^\dagger AUn\rangle = \sum_{k,l=1}^{N} U^\dagger_{mk}\, A_{kl}\, U_{ln} \tag{2.18}$$

2.2.2 Projecteurs et notation de Dirac

Enfin, nous ferons usage intensif des *projecteurs*. Soit $\mathcal{H}_1$ un sous-espace de $\mathcal{H}$ et $\mathcal{H}_2$ le sous-espace orthogonal. Tout vecteur $|\varphi\rangle$ se décompose de façon unique en un vecteur $|\varphi_1\rangle$ appartenant à $\mathcal{H}_1$ et un vecteur $|\varphi_2\rangle$ appartenant à $\mathcal{H}_2$

$$|\varphi\rangle = |\varphi_1\rangle + |\varphi_2\rangle, \quad |\varphi_1\rangle \in \mathcal{H}_1, \quad |\varphi_2\rangle \in \mathcal{H}_2, \quad \langle\varphi_1|\varphi_2\rangle = 0$$

On définit le projecteur $\mathcal{P}_1$ sur $\mathcal{H}_1$ par son action sur un vecteur arbitraire $|\varphi\rangle$

$$|\mathcal{P}_1\varphi\rangle = |\varphi_1\rangle \tag{2.19}$$

$\mathcal{P}_1$ est manifestement un opérateur linéaire, et c'est aussi un opérateur hermitien car si la décomposition de $|\chi\rangle$ en vecteurs appartenant à $\mathcal{H}_1$ et $\mathcal{H}_2$ est $|\chi\rangle = |\chi_1\rangle + |\chi_2\rangle$, alors

$$\begin{aligned}\langle\chi|\mathcal{P}_1\varphi\rangle &= \langle\chi|\varphi_1\rangle = \langle\chi_1|\varphi_1\rangle \\ \langle\chi|\mathcal{P}_1^\dagger\varphi\rangle &= \langle\mathcal{P}_1\chi|\varphi\rangle = \langle\chi_1|\varphi\rangle = \langle\chi_1|\varphi_1\rangle\end{aligned}$$

On remarque également que

$$|\mathcal{P}_1^2\varphi\rangle = |\mathcal{P}_1\varphi_1\rangle = |\varphi_1\rangle \Rightarrow \mathcal{P}_1^2 = \mathcal{P}_1$$

Inversement, tout opérateur linéaire qui vérifie $\mathcal{P}_1^\dagger\mathcal{P}_1 = \mathcal{P}_1$ est un projecteur.

Démonstration. On observe d'abord que $\mathcal{P}_1^\dagger = \mathcal{P}_1$, et ensuite que les vecteurs de la forme $|\mathcal{P}_1\varphi\rangle$ forment un sous-espace vectoriel $\mathcal{H}_1$ de $\mathcal{H}$. Si l'on écrit

$$|\varphi\rangle = |\mathcal{P}_1\varphi\rangle + (|\varphi\rangle - |\mathcal{P}_1\varphi\rangle) = |\mathcal{P}_1\varphi\rangle + |\varphi_2\rangle$$

alors $|\varphi_2\rangle$ est orthogonal à tout vecteur $|\mathcal{P}_1\chi\rangle$

$$\langle\varphi - \mathcal{P}_1\varphi|\mathcal{P}_1\chi\rangle = \langle\mathcal{P}_1\varphi - \mathcal{P}_1^2\varphi|\chi\rangle = 0$$

On a bien décomposé $|\varphi\rangle$ en $|\mathcal{P}_1\varphi\rangle$ et un vecteur du sous-espace orthogonal à $\mathcal{H}_1$.

La propriété $\mathcal{P}_1^2 = \mathcal{P}_1$ montre que les valeurs propres (voir le § 2.3.1) d'un projecteur sont 0 ou 1, et $\mathrm{Tr}\,\mathcal{P}_1$ (2.23) est égal à la dimension de l'espace de projection, comme on le voit aisément en écrivant $\mathcal{P}_1$ dans une base où il est diagonal : comme nous le verrons dans la section suivante, une telle base existe toujours car $\mathcal{P}_1$ est hermitien. De plus, on vérifiera dans l'exercice 2.4.6 les propriétés suivantes :

- Si $\mathcal{P}_1$ et $\mathcal{P}_1'$ sont des projecteurs sur $\mathcal{H}_1$ et $\mathcal{H}_1'$, respectivement, $\mathcal{P}_1\mathcal{P}_1'$ est un projecteur si et seulement si $\mathcal{P}_1\mathcal{P}_1' = \mathcal{P}_1'\mathcal{P}_1$. $\mathcal{P}_1\mathcal{P}_1'$ projette alors sur l'intersection $\mathcal{H}_1 \cap \mathcal{H}_1'$.
- $\mathcal{P}_1 + \mathcal{P}_1'$ est un projecteur si et seulement si $\mathcal{P}_1\mathcal{P}_1' = 0$. Dans ce cas, $\mathcal{H}_1$ et $\mathcal{H}_1'$ sont orthogonaux et $\mathcal{P}_1 + \mathcal{P}_1'$ projette sur la somme directe $\mathcal{H}_1 \oplus \mathcal{H}_1'$.
- Si $\mathcal{P}_1\mathcal{P}_1' = \mathcal{P}_1'\mathcal{P}_1$, alors $\mathcal{P}_1 + \mathcal{P}_1' - \mathcal{P}_1\mathcal{P}_1'$ projette sur l'union $\mathcal{H}_1 \cup \mathcal{H}_1'$. La seconde propriété est un cas particulier de celle-ci.

Notation de Dirac. Au lieu d'écrire $|A\varphi\rangle$, nous écrirons désormais $A|\varphi\rangle$ suivant une notation introduite par Dirac[4]. Le produit scalaire $\langle\chi|A\varphi\rangle$ s'écrira en notation de Dirac $\langle\chi|A|\varphi\rangle$. Les vecteurs $|\varphi\rangle$ de $\mathcal{H}$ sont appelés "kets", et les vecteurs $\langle\chi|$ de l'espace dual "bras". Le bra associé au ket $|\lambda\varphi\rangle$ est $\lambda^*\langle\varphi|$; en effet

$$\langle\lambda\varphi|\chi\rangle = \lambda^*\langle\varphi|\chi\rangle$$

Dans $\langle\chi|A|\varphi\rangle$, A agit à droite sur $|\varphi\rangle$: $\langle\chi|A|\varphi\rangle = \langle\chi|(A|\varphi\rangle)$, et non $\langle A\chi|\varphi\rangle$. Comme $(A|\varphi\rangle)^\dagger = \langle\varphi|A^\dagger$, il n'y a pas d'ambiguïté si A est hermitien. La notation de Dirac permet d'écrire très simplement les projecteurs. Soit $|\varphi\rangle$ un vecteur normalisé à l'unité : $\langle\varphi|\varphi\rangle = 1$. La décomposition de $|\chi\rangle$ suivant $|\varphi\rangle$ et un vecteur $|\chi_\perp\rangle$ orthogonal à $|\varphi\rangle$ est

$$|\chi\rangle = |\varphi\rangle\langle\varphi|\chi\rangle + (|\chi\rangle - |\varphi\rangle\langle\varphi|\chi\rangle) = |\varphi\rangle\langle\varphi|\chi\rangle + |\chi_\perp\rangle = \mathcal{P}_\varphi|\chi\rangle + |\chi_\perp\rangle$$

On peut donc écrire[5] :

$$\boxed{\mathcal{P}_\varphi = |\varphi\rangle\langle\varphi|} \tag{2.20}$$

4. Cette notation est commode et très largement utilisée, mais elle n'est pas exempte d'ambiguïtés. Elle est par exemple à éviter lorsque l'on traite du renversement du sens du temps, appendice A2.

5. Si $||\varphi||^2 \neq 1$, alors $P_\varphi = |\varphi\rangle\langle\varphi|/||\varphi||^2$.

Si une base orthonormée du sous-espace $\mathcal{H}_1$ est composée des vecteurs $\{|1\rangle, \ldots, |M\rangle\}$, $M \leq N$, alors $\mathcal{P}_1$ s'écrit :

$$\mathcal{P}_1 = \sum_{n=1}^{M} |n\rangle\langle n| \tag{2.21}$$

Si $M = N$, on obtient la décomposition de l'opérateur identité

$$\boxed{I = \sum_{n=1}^{N} |n\rangle\langle n|} \tag{2.22}$$

Cette relation est appelée *relation de fermeture*. Elle est souvent très commode dans les calculs. Par exemple, elle redonne simplement la loi de multiplication des matrices

$$(AB)_{nm} = \langle n|AB|m\rangle = \langle n|AIB|m\rangle = \sum_{l=1}^{N} \langle n|A|l\rangle\langle l|B|m\rangle = \sum_{l=1}^{N} A_{nl}B_{lm}$$

Donnons enfin une définition importante. La *trace* d'un opérateur est la somme de ses éléments diagonaux

$$\boxed{\operatorname{Tr} A = \sum_{n=1}^{N} A_{nn}} \tag{2.23}$$

Il est facile de montrer (exercice 2.4.2) que la trace est invariante dans un changement de base et que

$$\operatorname{Tr} AB = \operatorname{Tr} BA \tag{2.24}$$

2.3 Décomposition spectrale des opérateurs hermitiens

2.3.1 Diagonalisation d'un opérateur hermitien

Soit A un opérateur linéaire ; s'il existe un vecteur $|\varphi\rangle$ et un nombre complexe a tels que

$$A|\varphi\rangle = a|\varphi\rangle \tag{2.25}$$

alors $|\varphi\rangle$ est appelé *vecteur propre* et a *valeur propre* de A. On obtient les valeurs propres en résolvant l'équation en a

$$\det(A - aI) = 0 \tag{2.26}$$

Les vecteurs propres et valeurs propres des opérateurs hermitiens ont des propriétés remarquables.

Théorème. Les valeurs propres d'un opérateur hermitien sont réelles et les vecteurs propres correspondant à deux valeurs propres différentes sont orthogonaux.■

La démonstration est simple : il suffit de considérer le produit scalaire $\langle\varphi|A|\varphi\rangle$, où $|\varphi\rangle$ vérifie (2.25)

$$\begin{aligned}\langle\varphi|A|\varphi\rangle &= \langle\varphi|a\varphi\rangle = a||\varphi||^2 \\ &= \langle A\varphi|\varphi\rangle = \langle a\varphi|\varphi\rangle = a^*||\varphi||^2\end{aligned}$$

ce qui entraîne $a = a^*$; d'autre part si $A|\varphi\rangle = a|\varphi\rangle$ et $A|\chi\rangle = b|\chi\rangle$, alors

$$\langle\chi|A\varphi\rangle = a\langle\chi|\varphi\rangle = \langle A\chi|\varphi\rangle = b\langle\chi|\varphi\rangle$$

d'où $\langle\chi|\varphi\rangle = 0$ si $a \neq b$. Une conséquence immédiate de ce résultat est que les vecteurs propres d'un opérateur hermitien normalisés à l'unité forment une base orthonormée de $\mathcal{H}$ si les valeurs propres sont toutes distinctes, c'est-à-dire si les racines de l'équation (2.26) sont toutes distinctes. Cependant, il peut arriver que l'une (ou plusieurs) des racines de (2.26) soit racine multiple. Soit a_n une telle racine : la valeur propre a_n est alors dite *dégénérée*. Même dans ce cas, il est possible de former avec les vecteurs propres de A une base orthonormée de $\mathcal{H}$. En effet, on dispose du théorème suivant, que nous énonçons sans démonstration.

Théorème. Si un opérateur A est hermitien, il est toujours possible de trouver une matrice unitaire U (non unique) telle que $U^{-1}AU$ soit une matrice diagonale, dont les éléments diagonaux sont les valeurs propres qui apparaissent sur la diagonale un nombre de fois égal à leur dégénérescence

$$U^{-1}AU = \begin{pmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & \vdots \\ 0 & 0 & a_3 & 0 & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & a_N \end{pmatrix} \blacksquare \tag{2.27}$$

Soit a_n une valeur propre dégénérée, et soit $G(n)$ sa multiplicité dans (2.26) ; on dit aussi que a_n est $G(n)$ fois dégénérée. Il existe alors $G(n)$ vecteurs propres indépendants correspondant à cette valeur propre. Ces $G(n)$ vecteurs propres sous-tendent un sous-espace vectoriel de dimension $G(n)$, appelé *sous-espace de la valeur propre* a_n, où l'on peut trouver une base orthonormée (non unique) $|n,r\rangle,\ r = 1,\ \dots,\ G(n)$

$$A|n,r\rangle = a_n|n,r\rangle \tag{2.28}$$

Le projecteur $\mathcal{P}_n$ sur ce sous-espace vectoriel s'écrit d'après (2.21)

$$\mathcal{P}_n = \sum_{r=1}^{G(n)} |n, r\rangle\langle n, r| \tag{2.29}$$

La somme des $\mathcal{P}_n$ donne l'opérateur identité, puisque l'ensemble des vecteurs $|n, r\rangle$ forme une base de $\mathcal{H}$, et on obtient la relation de fermeture (2.22)

$$\boxed{\sum_n \mathcal{P}_n = \sum_n \sum_{r=1}^{G(n)} |n, r\rangle\langle n, r| = I} \tag{2.30}$$

Soit $|\varphi\rangle$ un vecteur quelconque de $\mathcal{H}$

$$A|\varphi\rangle = \sum_n A\mathcal{P}_n|\varphi\rangle = \sum_n a_n\mathcal{P}_n|\varphi\rangle$$

puisque $\mathcal{P}_n|\varphi\rangle$ appartient au sous-espace de la valeur propre a_n ; on peut donc faire l'identification suivante pour A

$$\boxed{A = \sum_n a_n\mathcal{P}_n = \sum_n \sum_{r=1}^{G(n)} |n, r\rangle a_n\langle n, r|} \tag{2.31}$$

Cette relation fondamentale est appelée *décomposition spectrale* de A. Réciproquement, un opérateur de la forme $\sum_n a_n\mathcal{P}_n$ est hermitien si $a_n = a_n^*$ et de valeurs propres a_n si $\mathcal{P}_n\mathcal{P}_m = \delta_{nm}\mathcal{P}_n$.

2.3.2 Diagonalisation d'une matrice 2×2 hermitienne

Nous aurons souvent l'occasion de diagonaliser des matrices 2×2 hermitiennes. La forme la plus générale d'une telle matrice dans une base $\{|1\rangle, |2\rangle\}$

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

est

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} a & b \\ b^* & a' \end{pmatrix}$$

où a et a' sont des nombres réels, b étant *a priori* complexe. Cependant, nous verrons qu'en mécanique quantique il est toujours possible de redéfinir la phase des vecteurs de base

$$|1\rangle \to |1'\rangle = \mathrm{e}^{\mathrm{i}\alpha}|1\rangle \qquad |2\rangle \to |2'\rangle = \mathrm{e}^{\mathrm{i}\beta}|2\rangle$$

Dans cette nouvelle base, l'élément de matrice A'_{12} de l'opérateur A est

$$A'_{12} = \langle 1'|A|2'\rangle = \mathrm{e}^{\mathrm{i}(\beta-\alpha)}\,\langle 1|A|2\rangle = \mathrm{e}^{\mathrm{i}(\beta-\alpha)}\,A_{12} = \mathrm{e}^{\mathrm{i}(\beta-\alpha)}\,b$$

Si $b = |b|\exp(\mathrm{i}\delta)$, il suffit de prendre $(\alpha - \beta) = \delta$ pour éliminer la phase de b, qui peut donc être choisi réel. Le cas le plus simple est celui où $a = a'$

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \tag{2.32}$$

Dans ce cas, on vérifie immédiatement que les deux vecteurs $|\chi_+\rangle$ et $|\chi_-\rangle$

$$|\chi_+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad |\chi_-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix} \tag{2.33}$$

sont vecteurs propres de A avec les valeurs propres $(a+b)$ et $(a-b)$, respectivement. Ce résultat très simple a une origine intéressante : soit U_P l'opérateur unitaire qui effectue une permutation des vecteurs de base $|1\rangle$ et $|2\rangle$

$$U_P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

L'opérateur U_P est de carré unité : $U_P^2 = I$, et ses valeurs propres sont donc ± 1. Les vecteurs propres correspondants sont $|\chi_+\rangle$ et $|\chi_-\rangle$. Mais on peut écrire A sous la forme

$$A = aI + bU_P$$

ce qui montre que A et U_P commutent : $AU_P = U_P A$ et, comme on le verra à la sous-section suivante, on peut alors trouver une base formée de vecteurs propres communs à A et U_P. La diagonalisation de A est simple parce que A commute avec une opération de symétrie, propriété que nous utiliserons souvent par la suite.

Dans le cas général $a \neq a'$, la propriété de symétrie n'est plus valable et la diagonalisation n'est pas aussi simple. Il est commode d'écrire A sous la forme

$$A = \begin{pmatrix} a+c & b \\ b & a-c \end{pmatrix} = aI + \sqrt{b^2+c^2}\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \tag{2.34}$$

où l'angle θ est défini par

$$\begin{aligned} c &= \sqrt{b^2+c^2}\,\cos\theta \\ b &= \sqrt{b^2+c^2}\,\sin\theta \end{aligned}$$

On notera que $\tan\theta = b/c$, et qu'il faut prendre garde à choisir la bonne détermination de θ. On vérifie alors que les vecteurs propres sont

$$|\chi_+\rangle = \begin{pmatrix} \cos\theta/2 \\ \sin\theta/2 \end{pmatrix} \qquad |\chi_-\rangle = \begin{pmatrix} -\sin\theta/2 \\ \cos\theta/2 \end{pmatrix} \tag{2.35}$$

correspondant aux valeurs propres $a + \sqrt{b^2 + c^2}$ et $a - \sqrt{b^2 + c^2}$, respectivement. On retrouve le cas précédent si $c = 0$, ce qui correspond à $\theta = \pm\pi/2$.

2.3.3 Ensemble complet d'opérateurs compatibles

Par définition, deux opérateurs A et B commutent si $AB = BA$, et dans ce cas leur *commutateur* $[A, B]$ défini par

$$[A, B] = AB - BA \tag{2.36}$$

est nul. Soit deux opérateurs hermitiens A et B qui commutent. On montre alors le théorème suivant.

Théorème. Soit A et B deux opérateurs hermitiens tels que $[A, B] = 0$. On peut alors trouver une base de $\mathcal{H}$ formée de vecteurs propres communs à A et B.■

Démonstration. Soit a_n les valeurs propres de A et $|n, r\rangle$ une base de $\mathcal{H}$ formée avec les vecteurs propres correspondants. Multiplions les deux membres de (2.28) par B et exploitons la commutation

$$BA|n, r\rangle = A(B|n, r\rangle) = a_n(B|n, r\rangle)$$

ce qui implique que le vecteur $B|n, r\rangle$ appartient au sous-espace de la valeur propre a_n. Si a_n est non dégénérée, ce sous-espace est de dimension un, $B|n, r\rangle$ est nécessairement proportionnel à $|n, r\rangle$ qui est donc également vecteur propre de B. Si a_n est dégénérée, nous pouvons seulement déduire que $B|n, r\rangle$ est nécessairement orthogonal à tout vecteur propre $|m, s\rangle$ de A, avec $m \neq n$

$$\langle m, s|B|n, r\rangle = \delta_{nm} B^{(n)}_{sr}$$

ce qui veut dire que dans la base $|n, r\rangle$ la matrice représentative de B est diagonale par blocs

$$B = \begin{pmatrix} B^{(1)} & 0 & 0 \\ 0 & B^{(2)} & 0 \\ 0 & 0 & B^{(3)} \end{pmatrix}$$

Chaque bloc $B^{(k)}$ peut être diagonalisé séparément par un changement de base affectant seulement chacun des sous-espaces, sans toucher à la diagonalisation de A puisqu'à l'intérieur de chaque sous-espace A est représenté par une matrice diagonale.

Réciproquement, supposons que l'on ait trouvé une base $|(n, p)r\rangle$ de $\mathcal{H}$ formée de vecteurs propres communs à A et B

$$A|(n, p)r\rangle = a_n|(n, p)r\rangle \qquad B|(n, p)r\rangle = b_p|(n, p)r\rangle$$

Il est alors évident que

$$[A, B]|(n, p)r\rangle = 0$$

et comme les vecteurs $|(n,p)r\rangle$ forment une base, $[A,B]=0$. Si $[A,B]=0$, il est possible que la donnée des valeurs propres a_n et b_p suffise à spécifier les vecteurs de base de façon unique, à une constante multiplicative de module un près ; il existe un vecteur $|(n,p)\rangle$ et un seul tel que

$$A|(n,p)\rangle = a_n|(n,p)\rangle \qquad B|(n,p)\rangle = b_p|(n,p)\rangle \tag{2.37}$$

On dira alors que A et B forment *un ensemble complet d'opérateurs compatibles.* S'il y a encore indétermination, c'est-à-dire s'il existe plusieurs vecteurs linéairement indépendants satisfaisant (2.37), il pourra arriver que la donnée des valeurs propres d'un troisième opérateur C commutant avec A et B lève l'indétermination. Un ensemble d'opérateurs hermitiens $A_1,\ldots,A_M$ commutant deux à deux et dont les valeurs propres définissent sans ambiguïté les vecteurs d'une base de $\mathcal{H}$ est appelé ensemble complet d'opérateurs compatibles (ou ensemble complet d'opérateurs qui commutent).

2.3.4 Opérateurs unitaires et opérateurs hermitiens

Les propriétés des opérateurs unitaires $U^\dagger = U^{-1}$ sont intimement liées à celles des opérateurs hermitiens, et en particulier ils peuvent toujours être diagonalisés. Le théorème de base pour les opérateurs unitaires s'énonce comme suit.

Théorème. (a) Les valeurs propres a_n d'un opérateur unitaire sont de module unité : $a_n = \exp(\mathrm{i}\alpha_n)$, α_n réel. (b) Les vecteurs propres correspondant à deux valeurs propres différentes sont orthogonaux. (c) La décomposition spectrale d'un opérateur unitaire s'écrit en fonction de projecteurs $\mathcal{P}_n$ sous la forme

$$U = \sum_n a_n\mathcal{P}_n = \sum_n \mathrm{e}^{\mathrm{i}\alpha_n}\mathcal{P}_n \quad \text{avec} \quad \sum_n \mathcal{P}_n = I \quad \blacksquare \tag{2.38}$$

La démonstration de (a) et (b) est triviale. Pour obtenir (c), on écrit :

$$U = \frac{1}{2}\,(U+U^\dagger) + \mathrm{i}\,\frac{1}{2\mathrm{i}}\,(U-U^\dagger) = A + \mathrm{i}B \tag{2.39}$$

Les opérateurs A et B sont hermitiens et $[A,B]=0$; (2.39) généralise aux opérateurs unitaires la décomposition en partie réelle et partie imaginaire d'un nombre complexe, les opérateurs hermitiens jouant le rôle des nombres réels. On peut diagonaliser simultanément A et B, et les vecteurs propres communs à A et B sont aussi vecteurs propres de U ; les valeurs propres de A et B sont $\cos\alpha_n$ et $\sin\alpha_n$, respectivement. L'opérateur C

$$C = \sum_n \alpha_n\mathcal{P}_n$$

est un opérateur hermitien et $U=\exp(\mathrm{i}C)$. Inversement, soit $A=\sum_n a_n\mathcal{P}_n$ un opérateur hermitien. L'opérateur

$$U = \sum_n \mathrm{e}^{\mathrm{i}\alpha a_n}\mathcal{P}_n = \mathrm{e}^{\mathrm{i}\alpha A} \tag{2.40}$$

est manifestement un opérateur unitaire. Cette écriture généralise aux opérateurs unitaires la représentation $\exp(\mathrm{i}\alpha)$ d'un nombre complexe de module un.

2.3.5 Fonctions d'un opérateur

En écrivant (2.40), nous avons introduit l'exponentielle d'un opérateur. Plus généralement, il est utile de savoir construire une fonction $f(A)$ d'un opérateur. Cette construction est immédiate si l'opérateur A peut être diagonalisé : $A = XDX^{-1}$, où D est une matrice diagonale dont les éléments diagonaux sont d_n. Supposons la fonction f définie par un développement de Taylor convergent dans un certain domaine du plan complexe $|z| < R$

$$f(z) = \sum_{p=0}^{\infty} c_p z^p$$

L'opérateur $f(A)$ sera donné par

$$f(A) = \sum_{p=0}^{\infty} c_p A^p = \sum_{p=0}^{\infty} c_p X D^p X^{-1} = X \left[\sum_{p=0}^{\infty} c_p D^p \right] X^{-1} \tag{2.41}$$

L'expression entre crochets n'est autre qu'une matrice diagonale d'éléments $f(d_n)$ bien définie si $|d_n| < R$ quel que soit n. En général, on pourra trouver un prolongement analytique pour $f(A)$ même si certaines valeurs propres d_n sont en dehors du rayon de convergence du développement de Taylor, de même que l'on prolonge analytiquement

$$\sum_{p=0}^{\infty} z^p = \frac{1}{1-z}$$

en dehors du rayon de convergence $|z| < 1$, pour toute valeur de z différente de un. Un cas particulièrement important est celui de l'exponentielle d'un opérateur

$$\exp A = \sum_{p=0}^{\infty} \frac{A^p}{p!} \tag{2.42}$$

Le rayon de convergence de ce développement étant infini, l'argument ci-dessus implique que $\exp A$ est bien défini par le développement (2.42) si A est diagonalisable (en fait, il n'est pas difficile de montrer que le développement est convergent dans tous les cas). Il faut prendre garde au fait qu'en général

$$\exp A \exp B \neq \exp B \exp A$$

une condition suffisante (mais non nécessaire !) pour l'égalité étant que A et B commutent (exercice 3.3.6).

En résumé, étant donné un opérateur hermitien A dont la décomposition spectrale est donnée par (2.31), il est immédiat de définir toute fonction de A par

$$f(A) = \sum_n f(a_n)\mathcal{P}_n \tag{2.43}$$

par exemple son exponentielle $\exp A$, son logarithme $\ln A$ ou sa *résolvante* $R(z,A)$

$$\mathrm{e}^{\,\mathrm{i}\alpha A} = \sum_n \mathrm{e}^{\,\mathrm{i}\alpha a_n}\,\mathcal{P}_n \tag{2.44}$$

$$\ln A = \sum_n (\ln a_n)\mathcal{P}_n \tag{2.45}$$

$$R(z,A) = (zI - A)^{-1} = \sum_n \frac{1}{z - a_n}\,\mathcal{P}_n \tag{2.46}$$

La résolvante $R(z,A)$ n'est bien sûr définie que si $z \neq a_n$ quel que soit n, et le logarithme si aucune des valeurs propres a_n n'est nulle.

2.4 Produit tensoriel de deux espaces vectoriels

2.4.1 Définition et propriétés du produit tensoriel

Dans cette section, nous allons construire l'espace produit tensoriel de deux espaces de Hilbert, construction que nous utiliserons épisodiquement aux chapitres 8 et 9, mais qui nous sera surtout indispensable[6] au chapitre 11. Soit $\mathcal{H}_A^{d_A}$ et $\mathcal{H}_B^{d_B}$ (ou simplement $\mathcal{H}_A$ et $\mathcal{H}_B$) deux espaces de Hilbert de dimensions respectives d_A et d_B, et choisissons dans ces espaces deux vecteurs $|\varphi_A\rangle \equiv |\varphi\rangle \in \mathcal{H}_A$ et $|\chi_B\rangle \equiv |\chi\rangle \in \mathcal{H}_B$. Le couple $\{|\varphi\rangle, |\chi\rangle\}$ peut être considéré comme un vecteur appartenant à un espace vectoriel de dimension $d_A d_B$, appelé *produit tensoriel* des espaces $\mathcal{H}_A^{d_A}$ et $\mathcal{H}_B^{d_B}$, noté $\mathcal{H}_A^{d_A} \otimes \mathcal{H}_B^{d_B}$ (ou simplement $\mathcal{H}_A \otimes \mathcal{H}_B$), que nous allons définir précisément ci-dessous.

Choisissons une base orthonormée $|i_A\rangle \equiv |i\rangle, i = 1\ldots, d_A$ de $\mathcal{H}_A$ et une base orthonormée $|m_B\rangle \equiv |m\rangle, m = 1,\ldots, d_B$ de $\mathcal{H}_B$. Les indices i, j, k étiquettent les vecteurs d'une base de $\mathcal{H}_A$ et m, n, p ceux d'une base de $\mathcal{H}_B$: afin d'éviter une prolifération d'indices, nous omettrons souvent les indices A et B et noterons $|i\rangle$ au lieu de $|i_A\rangle$, $|m\rangle$ au lieu de $|m_B\rangle$, etc., sauf en cas d'ambiguïté possible. Décomposons sur ces bases des vecteurs arbitraires $|\varphi\rangle \in \mathcal{H}_A$ et $|\chi\rangle \in \mathcal{H}_B$

$$|\varphi\rangle = \sum_{n=1}^{d_A} c_i\,|i\rangle \qquad |\chi\rangle = \sum_{m=1}^{d_B} d_m\,|m\rangle \tag{2.47}$$

6. Cette section peut être omise en première lecture. Elle sera indispensable pour aborder le chapitre 11.

L'espace $\mathcal{H}_A^{d_A} \otimes \mathcal{H}_B^{d_B}$ sera défini comme un espace à $d_A d_B$ dimensions où les couples $\{|i\rangle, |m\rangle\}$, notés $|i \otimes m\rangle$, ou $|i\rangle \otimes |m\rangle$, forment une base orthonormée

$$\langle j \otimes p \,| i \otimes m\rangle = \delta_{ji}\delta_{pm} \tag{2.48}$$

et le produit tensoriel des vecteurs $|\varphi\rangle$ et $|\chi\rangle$, noté $|\varphi \otimes \chi\rangle$, ou $|\varphi\rangle \otimes |\chi\rangle$, est le vecteur de composantes $c_i d_m$ dans cette base

$$|\varphi \otimes \chi\rangle = \sum_{i,m} c_i d_m |i \otimes m\rangle \tag{2.49}$$

On vérifie immédiatement la linéarité de l'opération produit tensoriel

$$\begin{aligned} |\varphi \otimes (\chi_1 + \lambda\chi_2)\rangle &= |\varphi \otimes \chi_1\rangle + \lambda|\varphi \otimes \chi_2\rangle \\ |(\varphi_1 + \lambda\varphi_2) \otimes \chi\rangle &= |\varphi_1 \otimes \chi\rangle + \lambda|\varphi_2 \otimes \chi\rangle \end{aligned} \tag{2.50}$$

Il faut également vérifier que la définition du produit tensoriel est indépendante du choix de la base. Soit $|\tilde{k}\rangle$ et $|\tilde{n}\rangle$ deux bases orthonormées de $\mathcal{H}_A$ et $\mathcal{H}_B$ déduites des bases $|i\rangle$ et $|m\rangle$ par des transformations unitaires respectives $R\,(R^{-1} = R^\dagger)$ et $S\,(S^{-1} = S^\dagger)$

$$|\tilde{k}\rangle = \sum_i R_{ki}\,|i\rangle \qquad |\tilde{p}\rangle = \sum_m S_{pm}\,|m\rangle$$

D'après (2.49), le produit tensoriel $|\tilde{k} \otimes \tilde{p}\rangle$ est donné par

$$|\tilde{k} \otimes \tilde{p}\rangle = \sum_{k,m} R_{ki} S_{pm}\,|i \otimes m\rangle$$

Par ailleurs, on peut écrire la décomposition de $|\varphi\rangle$ et $|\chi\rangle$ dans les bases respectives $|\tilde{k}\rangle$ et $|\tilde{p}\rangle$

$$|\varphi\rangle = \sum_{k=1}^{d_A} \tilde{c}_k\,|\tilde{k}\rangle \qquad |\chi\rangle = \sum_{p=1}^{d_B} \tilde{d}_p\,|\tilde{p}\rangle$$

Un calcul immédiat (exercice 2.5.12) montre que

$$\sum_{k,p} \tilde{c}_k \tilde{d}_p |\tilde{k} \otimes \tilde{p}\rangle = |\varphi \otimes \chi\rangle$$

où $|\varphi \otimes \chi\rangle$ est défini par (2.49). Le résultat pour $|\varphi \otimes \chi\rangle$ est donc indépendant du choix de la base.

Les vecteurs de $\mathcal{H}_A \otimes \mathcal{H}_B$ qui sont des produits tensoriels ne forment qu'un ensemble très limité (et même pas un sous-espace vectoriel !) des vecteurs de $\mathcal{H}_A \otimes \mathcal{H}_B$. Le vecteur le plus général est de la forme

$$|\Phi_{AB}\rangle = \sum_{i,m} b_{im}\,|i \otimes m\rangle \tag{2.51}$$

En général on ne pourra *pas* écrire le vecteur $|\Phi_{AB}\rangle$ comme un produit tensoriel $|\varphi \otimes \chi\rangle$. En effet, il faudrait que l'on puisse factoriser b_{im} sous la forme $c_i d_m$, ce qui est impossible sauf si les systèmes sont indépendants.

Le produit tensoriel $C = A \otimes B$ de deux opérateurs linéaires A et B agissant respectivement dans les espaces $\mathcal{H}_A$ et $\mathcal{H}_B$ est défini par son action sur le vecteur produit tensoriel $|\varphi \otimes \chi\rangle$

$$(A \otimes B)|\varphi \otimes \chi\rangle = |A\varphi \otimes B\chi\rangle \tag{2.52}$$

et ses éléments de matrice dans la base $|i \otimes m\rangle$ de $\mathcal{H}_A \otimes \mathcal{H}_B$ sont donc

$$\langle j \otimes p|A \otimes B|i \otimes m\rangle = A_{ji} B_{pm} \tag{2.53}$$

En général, un opérateur C agissant sur $\mathcal{H}_A \otimes \mathcal{H}_B$ n'est *pas* de la forme $A \otimes B$. Ses éléments de matrice sont

$$\langle j \otimes p|C|i \otimes m\rangle = C_{jp;im}$$

et sauf cas particulier on ne peut pas écrire $C_{jp;im}$ sous la forme factorisée $A_{ji} B_{pm}$. Deux cas particuliers intéressants de (2.52) sont $A = I_A$ et $B = I_B$, où I_A et I_B sont les opérateurs identité de $\mathcal{H}_A$ et de $\mathcal{H}_B$

$$(A \otimes I_B)|\varphi \otimes \chi\rangle = |A\varphi \otimes \chi\rangle \qquad (I_A \otimes B)|\varphi \otimes \chi\rangle = |\varphi \otimes B\chi\rangle \tag{2.54}$$

En termes d'éléments de matrice

$$\langle j \otimes p|A \otimes I_B|i \otimes m\rangle = A_{ji}\delta_{pm} \qquad \langle j \otimes p|I_A \otimes B|i \otimes m\rangle = \delta_{ji} B_{pm} \tag{2.55}$$

Enfin, si $|\varphi\rangle$ est vecteur propre de A avec la valeur propre a ($A|\varphi\rangle = a|\varphi\rangle$), alors $|\varphi \otimes \chi\rangle$ sera vecteur propre de $A \otimes I_B$ avec la valeur propre a

$$(A \otimes I_B)|\varphi \otimes \chi\rangle = a|\varphi \otimes \chi\rangle \tag{2.56}$$

On omet souvent d'écrire explicitement les opérateurs identité I_A et I_B, et une écriture courante pour (2.56) est

$$A|\varphi \otimes \chi\rangle = a|\varphi \otimes \chi\rangle \quad \text{ou simplement} \quad A|\varphi\chi\rangle = a|\varphi\chi\rangle \tag{2.57}$$

en supprimant le symbole du produit tensoriel. Comme la notation $\otimes$ est assez lourde, elle sera souvent omise sauf s'il y a une ambiguïté possible.

2.4.2 Espaces de dimension $d = 2$

Nous allons illustrer la notion de produit tensoriel en construisant l'espace produit tensoriel de deux espaces de dimension $d = 2$, $\mathcal{H}_A$ et $\mathcal{H}_B$. L'espace produit tensoriel $\mathcal{H}_{AB} \equiv \mathcal{H}_A \otimes \mathcal{H}_B$ est à quatre dimensions ($4 = 2 \times 2$). On choisit comme base orthonormée de $\mathcal{H}_A$ et de $\mathcal{H}_B$ les états de base $|+\rangle$ et $|-\rangle$

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.58}$$

qui seront notés génériquement $|\varepsilon\rangle, \varepsilon = \pm 1$. Suivant (2.49), les états de $\mathcal{H}_A \otimes \mathcal{H}_B$ se décomposent sur la base orthonormée $\varepsilon_A \otimes \varepsilon_B$

$$|+_A \otimes +_B\rangle, \quad |+_A \otimes -_B\rangle, \quad |-_A \otimes +_B\rangle, \quad |-_A \otimes -_B\rangle \tag{2.59}$$

Comme ci-dessus, on peut éliminer les indices A et B et écrire les vecteurs de base simplement

$$|+\otimes +\rangle, \quad |+\otimes -\rangle, \quad |-\otimes +\rangle, \quad |-\otimes -\rangle \tag{2.60}$$

Suivant (2.57), on omettra souvent la notation produit tensoriel en écrivant $|\varepsilon_A \varepsilon_B\rangle$ au lieu de $|\varepsilon_A \otimes \varepsilon_B\rangle$ ou bien $|++\rangle$ au lieu de $|+\otimes +\rangle$. Donnons un exemple de construction explicite d'un produit tensoriel : soit $|\varphi\rangle$ et $|\chi\rangle$ deux vecteurs arbitraires (normalisés) de $\mathcal{H}_A$ et de $\mathcal{H}_B$

$$\begin{aligned} |\varphi\rangle &= \lambda_A|+\rangle + \mu_A|-\rangle \qquad & |\lambda_A|^2 + |\mu_A|^2 = 1 \\ |\chi\rangle &= \lambda_B|+\rangle + \mu_B|-\rangle \qquad & |\lambda_B|^2 + |\mu_B|^2 = 1 \end{aligned}$$

Le produit tensoriel $|\varphi \otimes \chi\rangle$ est donné suivant (2.49) par

$$|\varphi\otimes\chi\rangle = \lambda_A\lambda_B|+\otimes+\rangle + \lambda_A\mu_B|+\otimes-\rangle + \mu_A\lambda_B|-\otimes+\rangle + \mu_A\mu_B|-\otimes-\rangle \tag{2.61}$$

Un vecteur générique $|\Psi\rangle \in \mathcal{H}_{AB}$ est de la forme

$$|\Psi\rangle = \alpha|+\otimes+\rangle + \beta|+\otimes-\rangle + \gamma|-\otimes+\rangle + \delta|-\otimes-\rangle \tag{2.62}$$

Les nombres $\alpha, \ldots, \delta$ sont arbitraires, en dehors de la condition de normalisation $|\alpha|^2 + \cdots + |\delta|^2 = 1$. Une condition nécessaire (et en fait suffisante) pour que $|\Psi\rangle$ puisse se mettre sous la forme (2.61) est que $\alpha\delta = \beta\gamma$, et *a priori* cette condition n'a aucune raison d'être valide. Un vecteur $|\Psi\rangle$ générique (2.62) n'est pas un produit tensoriel, et un vecteur qui n'est pas de la forme (2.61) est ce que nous appellerons au chapitre 11 un état intriqué. Un cas particulier important que nous utiliserons au chapitre 11 est l'état intriqué

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\Big(|+\otimes-\rangle - |-\otimes+\rangle\Big)$$

ou en notation allégée (2.57)

$$\boxed{|\Phi\rangle = \frac{1}{\sqrt{2}}\Big(|+-\rangle - |-+\rangle\Big)} \tag{2.63}$$

Cet état est manifestement intriqué car $\alpha = \delta = 0$ et $\beta = \gamma = 1/\sqrt{2}$: $\alpha\delta \neq \beta\gamma$. Il possède des propriétés de symétrie tout à fait remarquables : exercice 2.5.13.

2.5 Exercices

2.5.1 Produit scalaire et norme

1. Soit une norme $||\varphi||$ dérivant d'un produit scalaire : $||\varphi||^2 = (\varphi, \varphi)$. Montrer que cette norme vérifie l'inégalité triangulaire

$$||\chi + \varphi|| \leq ||\chi|| + ||\varphi||$$

ainsi que

$$\big| ||\chi|| - ||\varphi|| \big| \leq ||\chi + \varphi||$$

2. Vérifier également

$$||\chi + \varphi||^2 + ||\chi - \varphi||^2 = 2(||\chi||^2 + ||\varphi||^2)$$

Quelle est l'interprétation de cette égalité dans le plan réel $\mathbb{R}^2$? Inversement si une norme vérifie cette propriété dans un espace vectoriel *réel*, montrer que

$$(\varphi, \chi) = (\chi, \varphi) = \frac{1}{4}\left(||\chi + \varphi||^2 - ||\chi - \varphi||^2\right)$$

définit un produit scalaire. Ce produit scalaire doit vérifier

$$(\chi, \varphi_1 + \varphi_2) = (\chi, \varphi_1) + (\chi, \varphi_2) \qquad (\chi, \lambda\varphi) = \lambda(\chi, \varphi)$$

Dans le cas d'un espace vectoriel complexe, montrer que

$$(\chi, \varphi) = \frac{1}{4}\left[\left(||\chi + \varphi||^2 - ||\chi - \varphi||^2\right) - \mathrm{i}\left(||\chi + \mathrm{i}\varphi||^2 - ||\chi - \mathrm{i}\varphi||^2\right)\right]$$

2.5.2 Commutateurs et traces

1. Montrer que

$$[A, BC] = B[A, C] + [A, B]C \tag{2.64}$$

2. La trace (2.23) d'un opérateur est la somme des éléments diagonaux de sa matrice représentative dans une base donnée. Montrer que

$$\mathrm{Tr}\, AB = \mathrm{Tr}\, BA \tag{2.65}$$

et en déduire que la trace est invariante dans un changement de base $A \to A' = SAS^{-1}$. La trace d'un opérateur est (heureusement !) indépendante de la base.

3. Montrer que la trace est invariante par permutation circulaire

$$\mathrm{Tr}\, ABC = \mathrm{Tr}\, BCA = \mathrm{Tr}\, CAB \tag{2.66}$$

2.5.3 Déterminant et trace

1. Soit une matrice $A(t)$ dépendant d'un paramètre t vérifiant

$$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = A(t)\,B$$

Montrer que $A(t) = A(0)\exp(Bt)$. Quelle est la solution de

$$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = BA(t)\,?$$

2. Montrer que

$$\det \mathrm{e}^{At_1} \times \det \mathrm{e}^{At_2} = \det \mathrm{e}^{A(t_1+t_2)}$$

En déduire

$$\det \mathrm{e}^{A} = \mathrm{e}^{\mathrm{Tr}\,A}$$

ou de façon équivalente

$$\det B = \mathrm{e}^{\mathrm{Tr}\ \ln B} \tag{2.67}$$

Suggestion : obtenir une équation différentielle pour la fonction $g(t) = \det[\exp(At)]$. Les résultats sont évidents si A est diagonalisable.

2.5.4 Projecteur dans $\mathbb{R}^3$

Soit dans l'espace réel à trois dimensions $\mathbb{R}^3$ deux vecteurs $\vec{u}_1$ et $\vec{u}_2$ linéairement indépendants, mais non nécessairement orthogonaux et de norme quelconque, et $\mathcal{P}$ le projecteur sur le plan défini par ces deux vecteurs. Montrer que l'action de $\mathcal{P}$ sur un vecteur $\vec{V}$ s'écrit

$$\mathcal{P}\vec{V} = \sum_{i,j=1}^{2} C_{ij}^{-1}(\vec{V}\cdot\vec{u}_i)\vec{u}_j \tag{2.68}$$

où la matrice 2×2 $C_{ij} = \vec{u}_i \cdot \vec{u}_j$.

2. Généralisation : soit p vecteurs linéairement indépendants $\vec{u}_1, \ldots, \vec{u}_p$ dans $\mathbb{R}^N$, $p < N$. Écrire le projecteur sur l'espace vectoriel engendré par ces p vecteurs.

2.5.5 Théorème de la projection

Soit $\mathcal{H}_1$ un sous-espace vectoriel de $\mathcal{H}$ et $|\varphi\rangle \in \mathcal{H}$. Montrer qu'il existe alors un élément unique $|\varphi_1\rangle$ de $\mathcal{H}_1$ tel que la norme $||\varphi_1 - \varphi||$ soit minimale : $||\varphi_1 - \varphi||$ est la distance de $|\varphi\rangle$ à $\mathcal{H}_1$. Déterminer $|\varphi_1\rangle$.

2.5.6 Propriétés des projecteurs

1. Si $\mathcal{P}_1$ et $\mathcal{P}'_1$ sont des projecteurs sur $\mathcal{H}_1$ et $\mathcal{H}'_1$, respectivement, $\mathcal{P}_1\mathcal{P}'_1$ est un projecteur si et seulement si $\mathcal{P}_1\mathcal{P}'_1 = \mathcal{P}'_1\mathcal{P}_1$. $\mathcal{P}_1\mathcal{P}'_1$ projette alors sur l'intersection $\mathcal{H}_1 \cap \mathcal{H}'_1$.

2. $\mathcal{P}_1 + \mathcal{P}'_1$ est un projecteur si et seulement si $\mathcal{P}_1\mathcal{P}'_1 = 0$. Dans ce cas, $\mathcal{H}_1$ et $\mathcal{H}'_1$ sont orthogonaux et $\mathcal{P}_1 + \mathcal{P}'_1$ projette sur la somme directe $\mathcal{H}_1 \oplus \mathcal{H}'_1$.

3. Si $\mathcal{P}_1\mathcal{P}'_1 = \mathcal{P}'_1\mathcal{P}_1$, alors $\mathcal{P}_1 + \mathcal{P}'_1 - \mathcal{P}_1\mathcal{P}'_1$ projette sur l'union $\mathcal{H}_1 \cup \mathcal{H}'_1$. La propriété **2** est un cas particulier de ce résultat.

4. Soit un opérateur Ω tel que $\Omega^\dagger\Omega$ soit un projecteur

$$\Omega^\dagger\Omega = \mathcal{P}$$

Montrer que $\Omega\Omega^\dagger$ est aussi un projecteur. Suggestion : montrer que

$$\Omega|\varphi\rangle = 0 \Longleftrightarrow \mathcal{P}|\varphi\rangle = 0$$

2.5.7 Intégrale gaussienne

Soit A une matrice réelle $N \times N$ symétrique et strictement positive (*cf.* exercice 2.5.10). Montrer que l'intégrale multiple

$$I(b) = \int \prod_{i=1}^{N} \mathrm{d}x_i \exp\Big(-\frac{1}{2}\sum_{jk} x_j A_{jk} x_k + b_j x_j \Big)$$

vaut

$$I(b) = \frac{(2\pi)^{N/2}}{\sqrt{\det A}} \exp\Big(\frac{1}{2}\sum_{jk} b_j A^{-1}_{jk} b_k\Big) \tag{2.69}$$

Suggestion : écrire

$$\sum_{jk} x_j A_{jk} x_k = x^T A x = \langle x|A|x\rangle$$

où x est un vecteur colonne et x^T un vecteur ligne et effectuer le changement de variables

$$x' = x - A^{-1}b$$

Ces intégrales gaussiennes sont fondamentales en théorie des probabilités et interviennent dans nombre de problèmes de physique.

2.5.8 Commutateurs et valeur propre dégénérée

Soit trois matrices $N \times N$ A, B et C qui vérifient

$$[A, B] = 0 \qquad [A, C] = 0 \qquad [B, C] \neq 0$$

Montrer qu'au moins une valeur propre de A est dégénérée.

2.5.9 Matrices normales

Une matrice C est dite *normale* si elle commute avec la matrice hermitienne conjuguée

$$C^\dagger C = CC^\dagger$$

En écrivant :

$$C = \frac{1}{2}\,(C + C^\dagger) + \mathrm{i}\,\frac{1}{2\mathrm{i}}\,(C - C^\dagger) = A + \mathrm{i}B$$

montrer que C est diagonalisable.

2.5.10 Matrices positives

Une *matrice A* est dite *positive* si, quel que soit le vecteur $|\varphi\rangle \neq 0$, la valeur moyenne est réelle et positive : $\langle\varphi|A|\varphi\rangle \geq 0$. Elle est dite *strictement positive* si $\langle\varphi|A|\varphi\rangle > 0$.

1. Montrer que toute matrice positive est hermitienne et qu'une condition nécessaire et suffisante pour qu'une matrice soit positive est que ses valeurs propres soient toutes ≥ 0.

2. Montrer que dans un espace de Hilbert réel, où une matrice hermitienne est symétrique ($A = A^T$), une matrice positive n'est pas en général symétrique.

2.5.11 Identités opératorielles

1. Soit l'opérateur $f(t)$ fonction du paramètre t

$$f(t) = \mathrm{e}^{tA} B \mathrm{e}^{-tA}$$

où les opérateurs A et B sont représentés par des matrices $N \times N$. Montrer que

$$\frac{\mathrm{d}f}{\mathrm{d}t} = [A, f(t)] \qquad \frac{\mathrm{d}^2 f}{\mathrm{d}t^2} = [A, [A, f(t)]], \qquad \text{etc.}$$

En déduire

$$\mathrm{e}^{tA} B \mathrm{e}^{-tA} = B + \frac{t}{1!}\,[A, B] + \frac{t^2}{2!}\,[A, [A, B]] + \ldots \tag{2.70}$$

2. On suppose que A et B commutent tous deux avec leur commutateur $[A, B]$. Écrire une équation différentielle pour l'opérateur

$$g(t) = \mathrm{e}^{At}\,\mathrm{e}^{Bt}$$

et en déduire

$$\mathrm{e}^{A+B} = \mathrm{e}^{A}\,\mathrm{e}^{B}\,\mathrm{e}^{-\frac{1}{2}\,[A,B]} \tag{2.71}$$

Attention ! Cette identité n'est pas généralement valable. Elle n'est garantie que si $[A, [A, B]] = [B, [A, B]] = 0$. Montrer également avec les mêmes hypothèses

$$\mathrm{e}^{A}\,\mathrm{e}^{B} = \mathrm{e}^{B}\,\mathrm{e}^{A}\,\mathrm{e}^{[A,B]} \tag{2.72}$$

2.5.12 Indépendance du produit tensoriel par rapport au choix de la base

Vérifier que la définition (2.49) du produit tensoriel de deux vecteurs est bien indépendante du choix de la base dans $\mathcal{H}_A$ et $\mathcal{H}_B$.

2.5.13 Produit tensoriel de deux matrices 2×2

Écrire explicitement la matrice 4×4 $A \otimes B$ produit tensoriel des matrices 2×2 A et B

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

2.5.14 Propriétés de symétrie de $|\Phi\rangle$

Une propriété remarquable de $|\Phi\rangle$ (2.63) est son invariance par application des matrices du groupe $SL(2, \mathbb{C})$, groupe des matrices complexes 2×2 de déterminant unité (chapitre 19)

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \qquad \alpha\delta - \beta\gamma = 1$$

La transformation d'un vecteur $|\varphi\rangle$ s'écrit : $|\varphi\rangle \to |\varphi\rangle' = A|\varphi\rangle$. Les transformés de $|+\rangle$ et $|-\rangle$ sont

$$\begin{aligned} |+\rangle' &= A|+\rangle = \alpha|+\rangle + \gamma|-\rangle \\ |-\rangle' &= A|-\rangle = \beta|+\rangle + \delta|-\rangle \end{aligned}$$

Calculer $|+-\rangle' = (A \otimes A)|+-\rangle$ et $|-+\rangle' = (A \otimes A)|-+\rangle$ et montrer que le transformé de $|\Phi\rangle$ est

$$|\Phi\rangle' = \frac{1}{\sqrt{2}}\Big(|+-\rangle' - |-+\rangle'\Big) = (\alpha\delta - \beta\gamma)|\Phi\rangle = |\Phi\rangle$$

Un sous-groupe de $SL(2, \mathbb{C})$ est le groupe $SU(2)$ des matrices unitaires de dimension deux et de déterminant unité. Ce sous-groupe est relié de façon remarquable au groupe des rotations, et nous venons de montrer en particulier que $|\Phi\rangle$ est invariant par rotation.

2.6 Bibliographie

Le résultats sur les espaces vectoriels de dimension finie et les opérateurs se trouvent dans tout cours d'algèbre linéaire niveau première année de licence. Comme complément, on pourra consulter Isham [1995], chapitres 2 et 3 ou Nielsen et Chuang [2001], chapitre 2, où l'on trouvera une démonstration élégante du théorème de décomposition spectrale d'un opérateur hermitien.

Le produit tensoriel est traité dans Messiah [1959], chapitres VII et VIII ou Cohen-Tannoudji *et al.* [1973], compléments E_{III} et E_{IV}. Deux références plus récentes sont Isham [1995], chapitre 6 ou Basdevant et Dalibard [2001], annexe D.

Chapitre 3

Polarisation : photon et spin 1/2

Dans ce chapitre, nous allons mettre progressivement en place les concepts de base de la physique quantique à l'aide de deux exemples simples, en utilisant une approche heuristique, plus inductive que déductive. Nous partirons d'un phénomène familier, celui de la polarisation de la lumière, qui nous permettra d'introduire le formalisme mathématique nécessaire. Nous montrerons que la description de la polarisation conduit naturellement à faire appel à un espace vectoriel complexe à deux dimensions, et nous établirons la correspondance entre un état de polarisation et un vecteur de cet espace, appelé espace des états de polarisation. Nous passerons ensuite à la description quantique de la polarisation d'un photon et nous illustrerons la construction des amplitudes de probabilité comme produits scalaires dans cet espace. Le second exemple sera celui du spin 1/2, où l'espace des états est également de dimension deux. Nous construirons les états de spin 1/2 les plus généraux en utilisant l'invariance par rotation. Enfin, nous introduirons la dynamique, qui nous permettra de suivre l'évolution du vecteur d'état au cours du temps.

Alors que l'analogie avec la polarisation de la lumière nous servira de guide pour construire la théorie quantique de la polarisation d'un photon, nous ne disposerons pas d'une telle analogie classique pour construire celle du spin 1/2. Dans ce dernier cas, la construction de la théorie quantique sera faite *sans référence à une théorie classique*, à partir d'une hypothèse sur la dimension de l'espace des états et en nous appuyant sur des principes de symétrie.

3.1 Polarisation de la lumière et polarisation d'un photon

3.1.1 Polarisation d'une onde électromagnétique

La polarisation de la lumière – ou plus généralement d'une onde électromagnétique – est un phénomène bien connu lié au caractère vectoriel du

champ électromagnétique. Considérons une onde lumineuse plane monochromatique de fréquence ω se propageant dans le sens des z positifs. Le champ électrique $\vec{E}(t)$ en un point donné est un vecteur orthogonal à la direction de propagation. Il est donc situé dans le plan xOy et a pour composantes $\{E_x(t), E_y(t), E_z(t) = 0\}$ (figure 3.1). Le cas le plus général est celui d'une polarisation elliptique, où le champ électrique est de la forme

$$\vec{E}(t) = \begin{cases} E_x(t) = E_{0x}\cos(\omega t - \delta_x) \\ E_y(t) = E_{0y}\cos(\omega t - \delta_y) \end{cases} \tag{3.1}$$

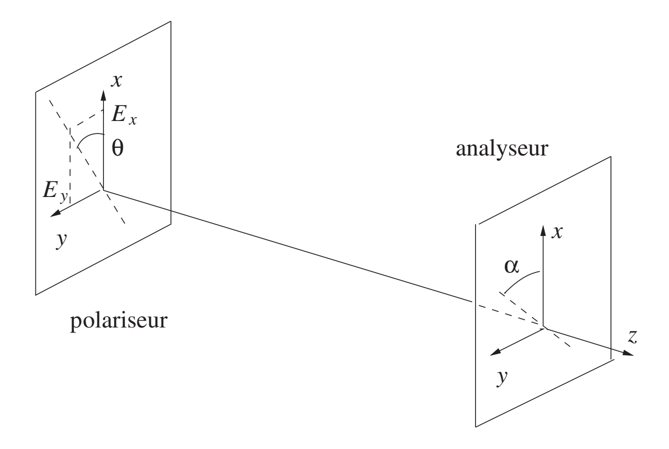


FIG. 3.1 – Ensemble polariseur-analyseur.

Nous n'avons pas explicité la dépendance en z car nous nous plaçons dans un plan $z =$cste. Par un changement d'origine des temps, il est toujours possible de choisir $\delta_x = 0,\ \delta_y = \delta$. L'intensité $\mathcal{I}$ de l'onde lumineuse est proportionnelle au carré du champ électrique

$$\mathcal{I} = \mathcal{I}_x + \mathcal{I}_y = k(E_{0x}^2 + E_{0y}^2) = kE_0^2 \tag{3.2}$$

où k est une constante de proportionnalité qu'il ne sera pas indispensable de préciser. Lorsque $\delta = 0$ ou π, la polarisation est *linéaire* : si l'on pose $E_{0x} = E_0\cos\theta$, $E_{0y} = E_0\sin\theta$, l'équation (3.1) pour $\delta_x = \delta_y = 0$ montre que le champ électrique vibre dans une direction $\hat{n}_\theta$ du plan xOy faisant un angle θ avec l'axe Ox. Une telle onde lumineuse s'obtient à l'aide d'un *polariseur* linéaire dont l'axe est parallèle à $\hat{n}_\theta$.

Lorsque nous nous intéressons uniquement à la polarisation de cette onde lumineuse, les paramètres pertinents sont les rapports $E_{0x}/E_0 = \cos\theta$ et $E_{0y}/E_0 = \sin\theta$, où l'on peut choisir θ dans l'intervalle $[0, \pi]$; E_0 est un simple facteur de proportionnalité qui ne joue aucun rôle dans la description de la polarisation. Nous pouvons faire correspondre aux ondes polarisées linéairement suivant Ox et Oy des vecteurs unitaires orthogonaux $|x\rangle$ et $|y\rangle$ du plan xOy

formant une base orthonormée de ce plan. À l'état de polarisation *linéaire* le plus général suivant $\hat{n}_\theta$ correspondra le vecteur $|\theta\rangle$ du plan xOy

$$|\theta\rangle = \cos\theta\,|x\rangle + \sin\theta\,|y\rangle \tag{3.3}$$

également de norme unité

$$\langle\theta|\theta\rangle = \cos^2\theta + \sin^2\theta = 1$$

La raison fondamentale qui conduit à utiliser un espace vectoriel pour décrire la polarisation est le *principe de superposition* : on peut décomposer un état de polarisation en deux (ou plusieurs) autres états, ou au contraire additionner vectoriellement deux états de polarisation. Pour illustrer la décomposition, faisons passer l'onde polarisée suivant $\hat{n}_\theta$ à travers un second polariseur, appelé *analyseur*, orienté suivant la direction $\hat{n}_\alpha$ du plan xOy faisant un angle α avec Ox (figure 3.1). Seule sera transmise la composante du champ électrique suivant $\hat{n}_\alpha$, sa projection sur $\hat{n}_\alpha$: l'amplitude du champ électrique sera multipliée par un facteur $\cos(\theta-\alpha)$ et l'intensité lumineuse à la sortie de l'analyseur sera réduite par un facteur $\cos^2(\theta-\alpha)$. Nous noterons $\overline{a}(\theta\to\alpha)$ le facteur de projection, que nous appellerons *amplitude de la polarisation* $\hat{n}_\theta$ *suivant* $\hat{n}_\alpha$, et nous remarquerons que cette amplitude n'est autre que le produit scalaire des vecteurs $|\theta\rangle$ et $|\alpha\rangle$

$$\overline{a}(\theta\to\alpha) = \langle\alpha|\theta\rangle = \cos(\theta-\alpha) = \hat{n}_\alpha\cdot\hat{n}_\theta \tag{3.4}$$

L'intensité à la sortie de l'analyseur est donnée par la loi de Malus

$$\mathcal{I} = \mathcal{I}_0|\overline{a}(\theta\to\alpha)|^2 = \mathcal{I}_0|\langle\alpha|\theta\rangle|^2 = \mathcal{I}_0\cos^2(\theta-\alpha) \tag{3.5}$$

si $\mathcal{I}_0$ est l'intensité à la sortie du polariseur. Une autre illustration de la décomposition est fournie par le dispositif de la figure 3.2a : à l'aide d'une lame biréfringente uniaxe perpendiculaire à la direction de propagation et dont l'axe optique se trouve dans le plan xOz, on décompose le faisceau lumineux en une onde polarisée suivant Ox et une onde polarisée suivant Oy. L'onde polarisée suivant Ox se propage dans une direction qui est celle du rayon extraordinaire, dévié à l'entrée et à la sortie de la lame, et celle polarisée suivant Oy suit le rayon ordinaire qui se propage en ligne droite. Dans les montages expérimentaux, on utilise souvent des lames biréfringentes sous la forme de *prismes polarisants*, ou *prismes de Wollaston*, qui divisent un faisceau incident en deux sous-faisceaux, l'un polarisé dans le plan de la figure et l'autre perpendiculairement à ce plan : figure 3.2b.

L'addition de deux états de polarisation est illustrée sur le dispositif de la figure 3.3 : les deux faisceaux sont recombinés par une seconde lame biréfringente symétrique de la première par rapport à un plan vertical avant de passer dans l'analyseur[1]. Afin de simplifier le raisonnement, nous négligeons

1. Cette recombinaison des amplitudes est possible parce que les deux faisceaux étant issus de la même source sont cohérents. Il serait bien sûr impossible d'additionner les amplitudes de deux faisceaux polarisés issus de sources différentes : le problème est identique à celui des interférences.

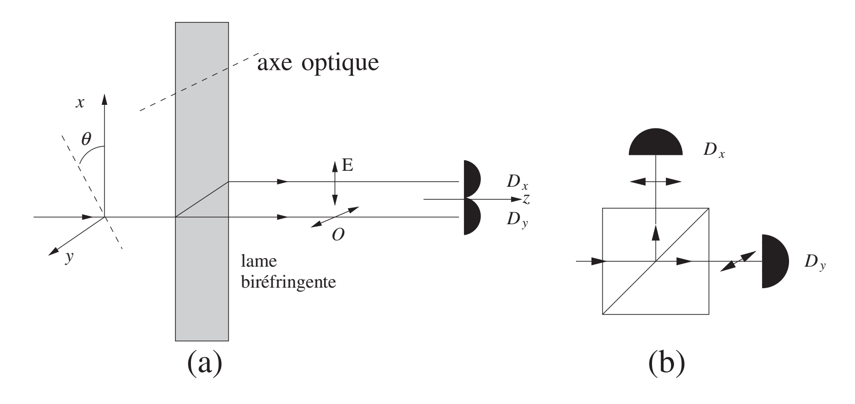


FIG. 3.2 – Décomposition de la polarisation par une lame biréfringente. (a) Le rayon ordinaire O est polarisé horizontalement, perpendiculairement au plan de la figure, le rayon extraordinaire E est polarisé verticalement, dans le plan de la figure. (b) Un prisme polarisant (ou de Wollaston) divise un faisceau incident en deux sous-faisceaux, l'un polarisé dans le plan de la figure (détecteur D_x) et l'autre polarisé perpendiculairement au plan de la figure (détecteur D_y).

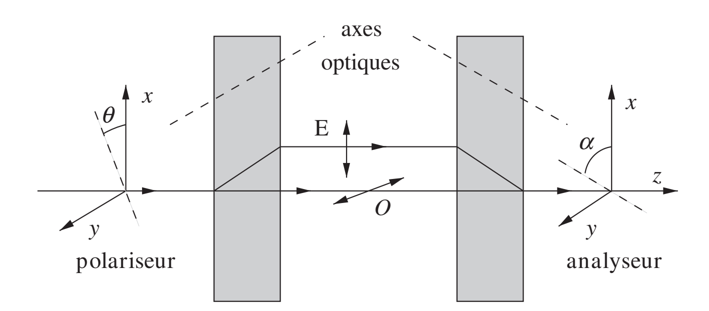


FIG. 3.3 – Décomposition et recombinaison de polarisations à l'aide de lames biréfringentes.

la différence de phase entre les deux ondes induite par la différence entre les indices ordinaire et extraordinaire dans les lames biréfringentes ; une façon de compenser cette différence d'indices est décrite dans l'exercice 3.3.5, mais elle conduirait à des complications dans la discussion qui va suivre. Si l'on néglige cette différence d'indices, l'onde lumineuse à la sortie de la seconde lame biréfringente est polarisée suivant $\hat{n}_\theta$: la recombinaison des deux faisceaux x et y donne la lumière initiale, polarisée suivant la direction $\hat{n}_\theta$, et l'intensité à la sortie de l'analyseur est réduite comme précédemment par le facteur $\cos^2(\theta - \alpha)$.

Si nous nous limitons à des états de polarisation linéaire, nous pouvons décrire tout état de polarisation comme un vecteur unitaire réel du plan xOy, dont une base orthonormée possible est formée des vecteurs $|x\rangle$ et $|y\rangle$. Mais si

nous voulons décrire une polarisation quelconque, nous devons introduire un espace *complexe* à deux dimensions $\mathcal{H}$. Cet espace sera *l'espace vectoriel des états de polarisation.* Revenons donc au cas général (3.1) en introduisant une notation complexe $\vec{\mathcal{E}} = (\mathcal{E}_x,\ \mathcal{E}_y)$ pour les amplitudes ondulatoires

$$\mathcal{E}_x = E_{0x}\,\mathrm{e}^{\mathrm{i}\delta_x} \qquad \mathcal{E}_y = E_{0y}\,\mathrm{e}^{\mathrm{i}\delta_y} \tag{3.6}$$

ce qui permet d'écrire (3.1) sous la forme

$$\begin{aligned} E_x(t) &= E_{0x}\cos(\omega t - \delta_x) = \mathrm{Re}\left(E_{0x}\,\mathrm{e}^{\mathrm{i}\delta_x}\mathrm{e}^{-\mathrm{i}\omega t}\right) = \mathrm{Re}\left(\mathcal{E}_x\,\mathrm{e}^{-\mathrm{i}\omega t}\right) \\ E_y(t) &= E_{0y}\cos(\omega t - \delta_y) = \mathrm{Re}\left(E_{0y}\,\mathrm{e}^{\mathrm{i}\delta_y}\mathrm{e}^{-\mathrm{i}\omega t}\right) = \mathrm{Re}\left(\mathcal{E}_y\,\mathrm{e}^{-\mathrm{i}\omega t}\right) \end{aligned} \tag{3.7}$$

Nous avons déjà remarqué qu'en raison de l'arbitraire sur l'origine des temps, seule la phase *relative* $\delta = (\delta_y - \delta_x)$ est physiquement pertinente et on peut multiplier simultanément $\mathcal{E}_x$ et $\mathcal{E}_y$ par un facteur de phase commun $\exp(\mathrm{i}\beta)$ sans conséquence physique. Il est toujours possible de choisir par exemple $\delta_x = 0$. L'intensité lumineuse est donnée par (3.2)

$$\mathcal{I} = k(|\mathcal{E}_x|^2 + |\mathcal{E}_y|^2) = k|\vec{\mathcal{E}}|^2 = kE_0^2 \tag{3.8}$$

Un cas particulier important de (3.7) est celui de la *polarisation circulaire*, où $E_{0x} = E_{0y} = E_0/\sqrt{2}$ et $\delta_y = \pm\pi/2$, lorsque l'on a choisi par convention $\delta_x = 0$. Si $\delta_y = +\pi/2$, l'extrémité du champ électrique décrit un cercle dans le plan xOy parcouru dans le sens trigonométrique. En effet $E_x(t)$ et $E_y(t)$ sont donnés par

$$\begin{aligned} E_x(t) &= \mathrm{Re}\left(\frac{E_0}{\sqrt{2}}\,\mathrm{e}^{-\mathrm{i}\omega t}\right) = \frac{E_0}{\sqrt{2}}\,\cos\omega t \\ E_y(t) &= \mathrm{Re}\left(\frac{E_0}{\sqrt{2}}\,\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{e}^{\mathrm{i}\pi/2}\right) = \frac{E_0}{\sqrt{2}}\,\cos(\omega t - \pi/2) = \frac{E_0}{\sqrt{2}}\,\sin\omega t \end{aligned} \tag{3.9}$$

Un observateur sur lequel arrive le rayon lumineux voit l'extrémité du vecteur champ électrique décrire dans le plan xOy un cercle de rayon $E_0/\sqrt{2}$ parcouru dans le sens trigonométrique : la polarisation correspondante est appelée *polarisation circulaire droite*[2]. Lorsque $\delta_y = -\pi/2$, on obtient une *polarisation circulaire gauche* : le cercle est parcouru dans le sens inverse du

2. Voir la figure 9.8. Notre définition des polarisations circulaires droite et gauche est celle adoptée en physique des particules élémentaires. Avec cette définition, la polarisation circulaire droite (gauche) correspond à une hélicité positive (négative), c'est-à-dire à une projection $+\hbar$ $(-\hbar)$ du spin du photon sur sa direction de propagation. Cependant, cette définition n'est pas universelle : les opticiens utilisent souvent la définition opposée, mais comme le remarque un opticien (E. Hecht [1987], chapitre 8) à propos de leur choix : "This choice of terminology is admittedly a bit awkward. Yet its use in optics is fairly common, even though it is completely antithetic to the more reasonable convention adopted in elementary particle physics."

sens trigonométrique

$$E_x(t) = \text{Re}\left(\frac{E_0}{\sqrt{2}}\,\text{e}^{-\text{i}\omega t}\right) = \frac{E_0}{\sqrt{2}}\cos\omega t$$
$$E_y(t) = \text{Re}\left(\frac{E_0}{\sqrt{2}}\,\text{e}^{-\text{i}\omega t}\text{e}^{-\text{i}\pi/2}\right) = \frac{E_0}{\sqrt{2}}\cos(\omega t + \pi/2) = -\frac{E_0}{\sqrt{2}}\sin\omega t \quad (3.10)$$

Ces états de polarisation circulaire droite et gauche sont obtenus expérimentalement en partant d'une polarisation linéaire à 45° par rapport aux axes et en déphasant de $\pm\pi/2$ le champ suivant Ox ou Oy par une lame quart d'onde.

En notation complexe, les champs $\mathcal{E}_x$ et $\mathcal{E}_y$ s'écrivent

$$\mathcal{E}_x = \frac{1}{\sqrt{2}}\,E_0 \qquad \mathcal{E}_y = \frac{1}{\sqrt{2}}\,E_0\,\text{e}^{\pm\text{i}\pi/2} = \frac{\pm\text{i}}{\sqrt{2}}\,E_0$$

où le signe (+) correspond à la polarisation circulaire droite et le signe (−) à la polarisation circulaire gauche. Le facteur de proportionnalité E_0 commun à $\mathcal{E}_x$ et $\mathcal{E}_y$ définit l'intensité de l'onde lumineuse et ne joue aucun rôle dans la description de la polarisation, qui est caractérisée par les vecteurs unitaires

$$\boxed{|D\rangle = -\frac{1}{\sqrt{2}}\,(|x\rangle + \text{i}|y\rangle) \qquad |G\rangle = \frac{1}{\sqrt{2}}\,(|x\rangle - \text{i}|y\rangle)} \quad (3.11)$$

Le signe (−) global dans la définition de $|D\rangle$ a été introduit par souci de cohérence avec les conventions du chapitre 9. L'équation (3.11) montre que la description mathématique de la polarisation nous amène naturellement à utiliser les vecteurs unitaires d'un espace vectoriel complexe bidimensionnel $\mathcal{H}$ dont les vecteurs $|x\rangle$ et $|y\rangle$ forment une base orthonormée possible.

Nous avons établi précédemment une correspondance entre une polarisation linéaire orientée suivant $\hat{n}_\theta$ et un vecteur unitaire $|\theta\rangle$ de $\mathcal{H}$, ainsi qu'une correspondance entre les deux polarisations circulaires et les deux vecteurs (3.11) de $\mathcal{H}$. Nous allons généraliser cette correspondance en construisant la polarisation correspondant au vecteur unitaire $|\Phi\rangle$ de $\mathcal{H}$ le plus général[3]

$$|\Phi\rangle = \lambda|x\rangle + \mu|y\rangle \qquad |\lambda|^2 + |\mu|^2 = 1 \quad (3.12)$$

Il est toujours possible de choisir λ réel (on vérifiera dans l'exercice 3.3.1 que la physique n'est pas modifiée si λ est complexe). Les nombres λ et μ peuvent alors être paramétrés par deux angles θ et η

$$\lambda = \cos\theta \qquad \mu = \sin\theta\,\text{e}^{\text{i}\eta}$$

Réalisons le dispositif suivant à l'aide de deux lames biréfringentes et d'un polariseur linéaire, sur lequel arrive une onde électromagnétique (3.7) : ce dispositif sera appelé *polariseur* (λ, μ).

3. Nous utilisons des lettres majuscules $|\Phi\rangle$ ou $|\Psi\rangle$ pour des vecteurs génériques de $\mathcal{H}$ de la forme (3.12) ou (3.16), afin qu'il n'y ait pas de confusion possible avec un angle, comme dans $|\theta\rangle$ ou $|\alpha\rangle$.

- Une première lame biréfringente déphase $\mathcal{E}_y$ de $-\eta$ en laissant $\mathcal{E}_x$ inchangé
$$\mathcal{E}_x \to \mathcal{E}_x^{(1)} = \mathcal{E}_x \qquad \mathcal{E}_y \to \mathcal{E}_y^{(1)} = \mathcal{E}_y\,\mathrm{e}^{-\mathrm{i}\eta}$$
- Le polariseur linéaire projette suivant $\hat{n}_\theta$
$$\begin{aligned}\vec{\mathcal{E}}^{(1)} \to \vec{\mathcal{E}}^{(2)} &= \left(\mathcal{E}_x^{(1)}\cos\theta + \mathcal{E}_y^{(1)}\sin\theta\right)\hat{n}_\theta \\ &= \left(\mathcal{E}_x\cos\theta + \mathcal{E}_y\sin\theta\,\mathrm{e}^{-\mathrm{i}\eta}\right)\hat{n}_\theta\end{aligned}$$
- La seconde lame biréfringente laisse $\mathcal{E}_x^{(2)}$ inchangé et déphase $\mathcal{E}_y^{(2)}$ de η
$$\mathcal{E}_x^{(2)} \to \mathcal{E}'_x = \mathcal{E}_x^{(2)} \qquad \mathcal{E}_y^{(2)} \to \mathcal{E}'_y = \mathcal{E}_y^{(2)}\mathrm{e}^{\mathrm{i}\eta}$$

La combinaison des trois opérations se traduit par la transformation $\vec{\mathcal{E}} \to \vec{\mathcal{E}}'$ de composantes

$$\begin{aligned}\mathcal{E}'_x &= \mathcal{E}_x\cos^2\theta + \mathcal{E}_y\sin\theta\cos\theta\,\mathrm{e}^{-\mathrm{i}\eta} = |\lambda|^2\mathcal{E}_x + \lambda\mu^*\mathcal{E}_y \\ \mathcal{E}'_y &= \mathcal{E}_x\sin\theta\cos\theta\,\mathrm{e}^{\mathrm{i}\eta} + \mathcal{E}_y\sin^2\theta = \lambda^*\mu\,\mathcal{E}_x + |\mu|^2\mathcal{E}_y\end{aligned} \tag{3.13}$$

L'opération (3.13) n'est autre que la *projection* sur $|\Phi\rangle$: en effet, si nous choisissons de représenter les vecteurs $|x\rangle$ et $|y\rangle$ par des vecteurs colonnes

$$|x\rangle = \begin{pmatrix}1\\0\end{pmatrix} \qquad |y\rangle = \begin{pmatrix}0\\1\end{pmatrix} \tag{3.14}$$

le projecteur $\mathcal{P}_\Phi$

$$\mathcal{P}_\Phi = |\Phi\rangle\langle\Phi| = \big(\lambda|x\rangle + \mu|y\rangle\big)\big(\lambda^*\langle x| + \mu^*\langle y|\big)$$

est représenté par la matrice

$$\mathcal{P}_\Phi = \begin{pmatrix}|\lambda|^2 & \lambda\mu^* \\ \lambda^*\mu & |\mu|^2\end{pmatrix} \tag{3.15}$$

Au champ incident $\vec{\mathcal{E}}$ (3.7), on peut faire correspondre un vecteur (non unitaire) $|\mathcal{E}\rangle$ de $\mathcal{H}$ de composantes complexes $\mathcal{E}_x$ et $\mathcal{E}_y$

$$|\mathcal{E}\rangle = \mathcal{E}_x|x\rangle + \mathcal{E}_y|y\rangle$$

À partir de $|\mathcal{E}\rangle$, on définit un vecteur unitaire $|\Psi\rangle$ par $|\mathcal{E}\rangle = E_0|\Psi\rangle$

$$|\Psi\rangle = \nu|x\rangle + \sigma|y\rangle \qquad |\nu|^2 + |\sigma|^2 = 1 \tag{3.16}$$

où

$$\nu = \frac{\mathcal{E}_x}{E_0} \qquad \sigma = \frac{\mathcal{E}_y}{E_0}$$

Le vecteur unitaire $|\Psi\rangle$ qui décrit la polarisation de l'onde (3.7) est appelé *vecteur de Jones.* D'après (3.13) et (3.15), le champ électrique sera à la sortie du polariseur (λ, μ)

$$|\mathcal{E}'\rangle = \mathcal{P}_\Phi|\mathcal{E}\rangle = E_0\mathcal{P}_\Phi|\Psi\rangle = E_0|\Phi\rangle\langle\Phi|\Psi\rangle \tag{3.17}$$

Nous venons de généraliser à un polariseur (λ, μ) ce que nous avions obtenu pour un polariseur linéaire : le polariseur (λ, μ) projette tout état de polarisation $|\Psi\rangle$ sur $|\Phi\rangle$ avec une amplitude égale à $\langle\Phi|\Psi\rangle$

$$\overline{a}(\Psi \to \Phi) = \langle\Phi|\Psi\rangle \tag{3.18}$$

À la sortie du polariseur, l'intensité est réduite par un facteur $|\overline{a}(\Psi \to \Phi)|^2 = |\langle\Phi|\Psi\rangle|^2$. Si l'état de polarisation est décrit par le vecteur unitaire $|\Phi\rangle$ (3.12), alors la transmission par le polariseur (λ, μ) se fait à 100 %. Au contraire, l'état de polarisation

$$|\Phi_\perp\rangle = -\mu^*|x\rangle + \lambda^*|y\rangle \tag{3.19}$$

est complètement arrêté par le polariseur (λ, μ). L'état de polarisation (3.16) est en général un état de polarisation elliptique. Il est facile de déterminer les caractéristiques de l'ellipse corrrespondante et le sens de parcours (exercice 3.3.1).

Les états $|\Phi\rangle$ et $|\Phi_\perp\rangle$ forment une base orthonormée de $\mathcal{H}$, obtenue à partir de la base $\{|x\rangle,\ |y\rangle\}$ par une transformation unitaire U

$$U = \begin{pmatrix} \lambda & \mu \\ -\mu^* & \lambda^* \end{pmatrix}$$

En résumé, nous avons montré qu'à un état de polarisation quelconque on peut faire corrrespondre un vecteur unitaire $|\Phi\rangle$ d'un espace à deux dimensions $\mathcal{H}$. Les vecteurs $|\Phi\rangle$ et $\exp(\mathrm{i}\beta)|\Phi\rangle$ représentent le même état de polarisation. En toute rigueur, on fait donc correspondre à un état de polarisation un vecteur à une phase près.

3.1.2 Polarisation d'un photon

Nous allons maintenant montrer que le formalisme mathématique utilisé ci-dessus pour décrire la polarisation d'une onde lumineuse se transpose sans modification à la description de la polarisation d'un photon. Cependant, cette identité du formalisme mathématique ne doit pas masquer que *l'interpétation physique subit une modification radicale.* Reprenons l'expérience de la figure 3.2 en remplaçant la source lumineuse par une source de photons uniques qui sont enregistrés individuellement par des photodétecteurs D_x et D_y. Ceux-ci détectent respectivement les photons polarisés suivant Ox et ceux polarisés suivant Oy. On observe alors

- que seul un des deux photodétecteurs est déclenché par un photon incident sur la lame. Comme les neutrons du chapitre 1, les photons arrivent entiers, ils ne se divisent jamais !
- que la probabilité p_x (p_y) de déclenchement de D_x (D_y) par un photon incident sur la lame est $\mathsf{p}_x = \cos^2\theta$ ($\mathsf{p}_y = \sin^2\theta$).

On doit nécessairement observer ce résultat si l'on veut retrouver l'optique classique à la limite où le nombre N de photons est grand[4] : en effet, si N_x et N_y sont les nombres de photons détectés par D_x et D_y, on doit avoir

$$\mathsf{p}_x = \lim_{N\to\infty} \frac{N_x}{N} \qquad \mathsf{p}_y = \lim_{N\to\infty} \frac{N_y}{N}$$

et $\mathcal{I}_x \propto N_x = N\cos^2\theta$, $\mathcal{I}_y \propto N_y = N\sin^2\theta$ à la limite où $N \to \infty$. Cependant, le sort *individuel* d'un photon ne peut pas être prédit : on connaît seulement sa *probabilité* de détection par D_x ou D_y. En physique quantique, les probabilités sont associées à des systèmes quantiques *individuels* alors qu'en physique classique les probabilités sont associées à des ensembles, et le recours aux probabilités est une façon de prendre en compte la complexité de phénomènes que nous ne pouvons pas (ou ne voulons pas) connaître dans le détail. Par exemple, dans le jeu de pile ou face, la connaissance parfaite des conditions initiales du lancer de la pièce, la prise en compte de la résistance de l'air, de la configuration du sol d'arrivée, etc. permettraient en théorie de prévoir le résultat. Quelques physiciens (dont de Broglie, Bohm, etc.) ont proposé que le caractère probabiliste de la mécanique quantique avait une origine analogue : si nous avions accès à des *variables additionnelles* pour le moment inconnues, alors nous pourrions prédire avec certitude le sort individuel de chaque photon. Cette hypothèse de variables additionnelles est utile lorsque l'on examine des fondements de la physique quantique. Toutefois, nous verrons au chapitre 11 que, à moins de supposer des interactions à distance instantanées, de telles variables sont exclues par l'expérience. En résumé, la connaissance de l'état de polarisation d'un photon permet de déterminer la probabilité qu'il soit transmis par un analyseur : cette probabilité est attachée individuellement à chaque photon, mais pour vérifier cette loi de probabilité, il faut effectuer un grand nombre d'expériences sur des photons tous dans le même état. En théorie classique des probabilités, chaque individu de l'ensemble est dans un état où tous ses paramètres sont déterminés, même s'ils nous sont inconnus, mais ce n'est pas le cas en physique quantique. Nous aurons l'occasion de revenir sur cette question au chapitre 11.

Cependant, la seule donnée de probabilités ne fournit qu'une description très incomplète de la polarisation d'un photon. Une description complète requiert l'introduction d'*amplitudes de probabilité*, et non simplement de probabilités. Les amplitudes de probabilité, notées a (nous soulignons la différence

4. Cet énoncé est correct, mais nous passons pour le moment sous silence des problèmes qui seront examinés dans les sections 17.2 et 17.4 : certains effets physiques sont différents, selon que l'on a affaire à N photons isolés arrivant un par un, à un état cohérent contenant une moyenne de N photons ou à un état de Fock à N photons.

entre les amplitudes ondulatoires de la sous-section précédente et les amplitudes de probabilité en utilisant une notation différente : a au lieu de $\overline{a}$), sont des nombres complexes, et les probabilités sont données par leur module carré $|a|^2$. Pour mettre en évidence le caractère incomplet de la seule donnée des probabilités, reprenons le dispositif de la figure 3.3. Entre les deux lames, un photon suit soit le trajet du rayon extraordinaire polarisé suivant Ox, étiqueté “trajet x”, soit le trajet du rayon ordinaire polarisé suivant Oy, étiqueté “trajet y”. Dans un raisonnement purement probabiliste, un photon suivant le trajet x aurait une probabilité $\cos^2\theta\cos^2\alpha$ d'être transmis par l'analyseur, et le photon suivant le trajet y une probabilité $\sin^2\theta\sin^2\alpha$, ce qui donnerait une probabilité totale

$$\mathsf{p}_{\rm tot} = \cos^2\theta\cos^2\alpha + \sin^2\theta\sin^2\alpha \tag{3.20}$$

qu'un photon soit transmis par l'analyseur. Ce n'est pas ce que donne l'expérience, qui confirme le résultat établi précédemment par un raisonnement ondulatoire

$$\mathsf{p}_{\rm tot} = \cos^2(\theta-\alpha)$$

Il faut raisonner en amplitudes de probabilité, comme nous l'avons fait pour l'amplitude d'une onde : les amplitudes de probabilité obéissent aux mêmes règles que les amplitudes ondulatoires, ce qui garantit que les résultats de l'optique sont reproduits lorsque le nombre de photons $N\to\infty$. L'amplitude de probabilité pour qu'un photon linéairement polarisé suivant la direction $\hat{n}_\theta$ soit polarisé suivant la direction $\hat{n}_\alpha$ est donnée par (3.4) : $a(\theta\to\alpha) = \cos(\theta-\alpha) = \hat{n}_\theta\cdot\hat{n}_\alpha$. On obtient le tableau suivant pour les amplitudes de probabilité intervenant dans l'expérience de la figure 3.3

$$\begin{array}{lcl} a(\theta\to x) & = & \cos\theta \qquad a(x\to\alpha)=\cos\alpha \\ a(\theta\to y) & = & \sin\theta \qquad a(y\to\alpha)=\sin\alpha \end{array}$$

Cet exemple permet d'illustrer les règles qui régissent les combinaisons d'amplitudes de probabilité. L'amplitude de probabilité a_x pour que le photon incident suivant le trajet x soit transmis par l'analyseur est

$$a_x = a(\theta\to x)a(x\to\alpha) = \cos\theta\cos\alpha$$

Cette expression met en évidence la règle de *factorisation* des amplitudes : a_x est le produit des amplitudes $a(\theta\to x)$ et $a(x\to\alpha)$. Cette règle de factorisation garantit que la règle correspondante pour les probabilités est bien vérifiée. On a de même

$$a_y = a(\theta\to y)a(y\to\alpha) = \sin\theta\sin\alpha$$

Si la configuration de l'expérience ne permet pas de savoir quel trajet a suivi le photon, alors on doit ajouter les amplitudes. L'amplitude de probabilité totale pour que le photon soit transmis par l'analyseur est donc

$$a_{\rm tot} = a_x + a_y = \cos\theta\cos\alpha + \sin\theta\sin\alpha = \cos(\theta-\alpha) \tag{3.21}$$

et la probabilité correspondante $\cos^2(\theta - \alpha)$, en accord avec le résultat (3.5) de l'optique classique. S'il existe une possibilité de distinguer entre les deux trajets, alors les interférences sont détruites et il faut ajouter les probabilités suivant (3.20).

Les règles de combinaison des amplitudes de probabilité étant les mêmes que pour les amplitudes ondulatoires, ces règles seront satisfaites si l'on décrit l'état de polarisation d'un photon par un vecteur unitaire dans un espace vectoriel à deux dimensions $\mathcal{H}$, appelé *espace des états*, dans le cas présent l'espace des états de polarisation. Lorsqu'un photon est polarisé linéairement suivant Ox (Oy), nous ferons correspondre à son état de polarisation un vecteur $|x\rangle$ ($|y\rangle$) de cet espace. Un tel état de polarisation est obtenu en faisant passer le photon à travers un polariseur linéaire orienté suivant Ox (Oy). La probabilité qu'un photon polarisé suivant Ox soit transmis par un analyseur orienté suivant Oy est nulle : l'amplitude de probabilité $a(x \to y) = 0$. Inversement, la probabilité qu'un photon polarisé suivant Ox ou Oy soit transmis par un analyseur dans la même direction est égale à un

$$|a(x \to x)| = |a(y \to y)| = 1 \qquad a(x \to y) = a(y \to x) = 0$$

Ces relations sont satisfaites si $|x\rangle$ et $|y\rangle$ forment une base orthonormée de $\mathcal{H}$ et si nous identifions les amplitudes de probabilité aux produits scalaires

$$a(x \to x) = \langle x|x\rangle = 1 \quad a(y \to y) = \langle y|y\rangle = 1 \quad a(y \to x) = \langle x|y\rangle = 0 \quad (3.22)$$

L'état de polarisation *linéaire* le plus général est un état dont la polarisation fait un angle θ avec Ox ; cet état sera représenté par le vecteur

$$|\theta\rangle = \cos\theta\,|x\rangle + \sin\theta\,|y\rangle \qquad (3.23)$$

Les équations (3.22) et (3.23) assurent que les amplitudes de probabilité écrites précédemment sont correctement données par les produits scalaires, par exemple

$$a(\theta \to x) = \langle x|\theta\rangle = \cos\theta$$

ou en général, si $|\alpha\rangle$ est un état de polarisation linéaire

$$a(\theta \to \alpha) = \langle \alpha|\theta\rangle = \cos(\theta - \alpha)$$

L'état de polarisation le plus général sera décrit par un vecteur unitaire, appelé *vecteur d'état*

$$|\Phi\rangle = \lambda|x\rangle + \mu|y\rangle \qquad |\lambda|^2 + |\mu|^2 = 1$$

Comme dans le cas ondulatoire, les vecteurs $|\Phi\rangle$ et $\exp(\mathrm{i}\beta)|\Phi\rangle$ représentent le même état physique : un état physique est représenté par un vecteur à une phase près dans l'espace des états. *L'amplitude de probabilité pour trouver un état de polarisation* $|\Psi\rangle$ *dans* $|\Phi\rangle$ sera donnée par le produit scalaire $\langle\Phi|\Psi\rangle$, et la projection sur un état de polarisation déterminé sera réalisée par le

dispositif décrit à la sous-section précédente. En résumé, nous avons illustré sur un exemple concret, celui de la polarisation d'un photon, la construction de l'espace de Hilbert des états de polarisation.

La polarisation d'un photon est un exemple de *propriété physique quantique*. L'interprétation d'une propriété physique quantique diffère radicalement de celle d'une propriété physique classique. Nous allons l'illustrer en examinant la polarisation d'un photon. Nous nous limiterons dans un premier temps au cas le plus simple des états de polarisation linéaire. À l'aide d'un polariseur linéaire orienté suivant Ox, préparons un ensemble de photons tous dans un état de polarisation linéaire $|x\rangle$: les photons arrivent un par un sur le polariseur, et seuls ceux dont la polarisation est parallèle à Ox sont transmis. C'est la phase de *préparation du système quantique*, où l'on ne conserve donc que les photons ayant traversé le polariseur orienté suivant Ox. La phase suivante, ou *phase de test*, consiste à tester cette polarisation en faisant passer le photon dans un analyseur linéaire : si cet analyseur est parallèle à Ox, les photons sont transmis avec une probabilité un, s'il est parallèle à Oy avec une probabilité nulle. Dans les deux cas, le résultat du test peut être prédit avec certitude. La propriété physique "polarisation d'un photon préparé dans l'état $|x\rangle$" prend des valeurs *certaines* si l'on choisit la base $\{|x\rangle, |y\rangle\}$ pour le test.

En revanche, si nous utilisons des analyseurs orientés dans la direction $\hat{n}_\theta$, correspondant à l'état $|\theta\rangle$ (3.23), et dans la direction perpendiculaire $\hat{n}_{\theta_\perp}$, correspondant à l'état

$$|\theta_\perp\rangle = -\sin\theta\,|x\rangle + \cos\theta\,|y\rangle \tag{3.24}$$

nous pouvons seulement prédire une *probabilité* de transmission $|\langle\theta|x\rangle|^2 = \cos^2\theta$ dans le premier cas et $|\langle\theta_\perp|x\rangle|^2 = \sin^2\theta$ dans le second. La propriété physique "polarisation du photon dans l'état $|x\rangle$" n'a pas de valeur certaine dans la base $\{|\theta\rangle, |\theta_\perp\rangle\}$. Autrement dit, la propriété physique polarisation est attachée à une base déterminée, et les deux bases $\{|x\rangle, |y\rangle\}$ et $\{|\theta\rangle, |\theta_\perp\rangle\}$ sont dites *incompatibles* (sauf pour $\theta = 0$ et $\theta = \pi/2$). Un cas intéressant de bases incompatibles est celui des *bases complémentaires* : dans un espace de Hilbert de dimension N, deux bases $|n\rangle$ et $|\alpha\rangle$ sont complémentaires si $|\langle\alpha|n\rangle|^2 = 1/N$, quels que soient $|n\rangle$ et $|\alpha\rangle$. Les bases de polarisation linéaires $\{|x\rangle, |y\rangle\}$ et $\{|\theta\rangle, |\theta_\perp\rangle\}$ pour $\theta = \pi/4$ sont complémentaires ; de même une base de polarisation circulaire $\{|D\rangle, |G\rangle\}$ (3.11) est complémentaire de toute base de polarisation linéaire $\{|\theta\rangle, |\theta_\perp\rangle\}$.

La discussion précédente mérite d'être précisée sur deux points. Tout d'abord, il est clair que l'on ne peut pas déterminer la polarisation d'un photon *isolé*. Le test de polarisation suppose que l'on dispose d'un nombre $N \gg 1$ de photons préparés dans des conditions identiques. Supposons que l'on prépare N photons dans un certain état de polarisation et qu'on les teste en orientant un analyseur linéaire suivant Ox ; si on constate – dans la limite des imperfections du dispositif expérimental – que les photons traversent l'analyseur avec une probabilité de 100 %, on pourra en déduire que les photons

ont été préparés dans l'état $|x\rangle$. L'observation d'un seul photon ne permet évidemment pas d'arriver à cette conclusion, sauf si on connaît par avance la base dans laquelle il a été préparé. Le second point est que si les photons sont transmis avec une probabilité $\cos^2\theta$, on ne pourra pas en déduire qu'ils ont été préparés dans l'état de polarisation linéaire (3.23). En effet, on observera la même probabilité de transmission si les photons ont été préparés dans l'état de polarisation elliptique (3.12), avec

$$\lambda = \cos\theta\, \mathrm{e}^{\,\mathrm{i}\delta_x} \qquad\qquad \mu = \sin\theta\, \mathrm{e}^{\,\mathrm{i}\delta_y}$$

et il faut effectuer une autre série d'expériences avec une orientation différente de l'analyseur pour mesurer les phases (exercice 3.3.1). Seul un test dont les résultats ont une probabilité 0 ou 1 permet de déterminer sans ambiguïté l'état de polarisation des photons.

Dans la représentation (3.14) des vecteurs de base de $\mathcal{H}$, les projecteurs $\mathcal{P}_x$ et $\mathcal{P}_y$ sur les états $|x\rangle$ et $|y\rangle$ sont représentés par les matrices

$$\mathcal{P}_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad\qquad \mathcal{P}_y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

qui commutent : $[\mathcal{P}_x, \mathcal{P}_y] = 0$. Les deux opérateurs sont compatibles suivant la définition du § 2.3.3. Les projecteurs $\mathcal{P}_\theta$ et $\mathcal{P}_{\theta_\perp}$, que l'on calcule immédiatement à partir de (3.15), sont donnés par

$$\mathcal{P}_\theta = \begin{pmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{pmatrix} \qquad\qquad \mathcal{P}_{\theta_\perp} = \begin{pmatrix} \sin^2\theta & -\sin\theta\cos\theta \\ -\sin\theta\cos\theta & \cos^2\theta \end{pmatrix}$$

Ils commutent entre eux, mais ne commutent ni avec $\mathcal{P}_x$, ni avec $\mathcal{P}_y$: $\mathcal{P}_x$ et $\mathcal{P}_\theta$ par exemple sont incompatibles. La commutation (ou la non commutation) d'opérateurs traduit mathématiquement la compatibilité (ou la non compatibilité) de propriétés physiques.

Un autre choix de base consiste à utiliser les états de polarisation circulaire droite $|D\rangle$ et gauche $|G\rangle$ (3.11). La base $\{|D\rangle, |G\rangle\}$ est incompatible avec toute base formée avec des états de polarisation linéaire (et même complémentaire de ces bases). Les projecteurs $\mathcal{P}_D$ et $\mathcal{P}_G$ sur les états de polarisation circulaire sont

$$\mathcal{P}_D = \frac{1}{2}\begin{pmatrix} 1 & -\mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix} \qquad\qquad \mathcal{P}_G = \frac{1}{2}\begin{pmatrix} 1 & \mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix} \tag{3.25}$$

Avec $\mathcal{P}_D$ et $\mathcal{P}_G$ on forme l'opérateur hermitique remarquable Σ_z

$$\boxed{\Sigma_z = \mathcal{P}_D - \mathcal{P}_G = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}} \tag{3.26}$$

Cet opérateur a pour vecteurs propres les états $|D\rangle$ et $|G\rangle$ dont les valeurs propres respectives sont $+1$ et -1

$$\Sigma_z|D\rangle = |D\rangle \qquad\qquad \Sigma_z|G\rangle = -|G\rangle \tag{3.27}$$

Ce résultat suggère d'associer à la propriété physique “polarisation circulaire” un opérateur hermitien Σ_z dont les vecteurs propres sont $|D\rangle$ et $|G\rangle$. Nous verrons au chapitre 9 que $\hbar\Sigma_z = J_z$ est l'opérateur représentant la propriété physique composante suivant Oz du moment angulaire (ou spin) du photon. Nous verrons également que $\exp(-\mathrm{i}\theta\Sigma_z)$ est l'opérateur qui effectue des rotations d'un angle θ autour de l'axe Oz. Un calcul simple (exercice 3.3.3) donne en effet

$$\exp(-\mathrm{i}\theta\Sigma_z) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{3.28}$$

et $\exp(-\mathrm{i}\theta\Sigma_z)$ transforme bien l'état $|x\rangle$ en l'état $|\theta\rangle$ et $|y\rangle$ en $|\theta_\perp\rangle$

$$\exp(-\mathrm{i}\theta\Sigma_z)|x\rangle = |\theta\rangle \qquad \exp(-\mathrm{i}\theta\Sigma_z)|y\rangle = |\theta_\perp\rangle \tag{3.29}$$

3.1.3 Cryptographie quantique

La cryptographie quantique est une invention récente fondée sur l'incompatibilité de deux bases différentes d'états de polarisation linéaire. La cryptographie usuelle repose sur une clé de chiffrage connue seulement de l'expéditeur et du destinataire. Ce système est appelé à *clé secrète*. Il est en principe très sûr[5], mais il faut que l'expéditeur et le destinataire aient le moyen de se transmettre la clé sans que celle-ci soit interceptée par un espion. Or la clé doit être changée fréquemment, car une suite de messages codés avec la même clé est susceptible de révéler des régularités permettant le déchiffrage du message par une tierce personne. Le processus de transmission d'une clé secrète est un processus à risque, et c'est pour cette raison que l'on préfère maintenant les systèmes fondés sur un principe différent, dits systèmes à *clé publique*, où la clé est diffusée publiquement, par exemple sur Internet. Le système à clé publique courant[6] est fondé sur la difficulté de décomposer un nombre très grand N en facteurs premiers, alors que l'opération inverse est immédiate : sans calculette, on obtiendra en quelques secondes $137 \times 53 = 7261$, mais étant donné 7261, cela prendra un certain temps pour le décomposer en facteurs premiers. Avec les meilleurs algorithmes actuels, le temps de calcul sur ordinateur nécessaire pour décomposer un nombre N en facteurs premiers croît avec N comme $\simeq \exp[1.9(\ln N)^{1/3}(\ln\ln N)^{2/3}]$. Le record actuel (2010) est de 232 chiffres, et il faut plusieurs mois à une grappe de PC pour arriver au résultat. Dans le système de chiffrage à clé publique, le destinataire, appelé conventionnellement Bob, diffuse publiquement à l'expéditeur, appelé conventionnellement Alice, un nombre très grand $N = pq$ produit de deux nombres premiers p et q. Ce nombre suffit à Alice pour chiffrer le message, mais il faut disposer des nombres p et q pour le déchiffrer. Bien sûr, un espion (appelé par convention

5. Un chiffrage absolument sûr a été découvert par Vernam en 1917. Cependant, la sécurité absolue suppose que la clé soit aussi longue que le message et ne soit utilisée qu'une seule fois !

6. Appelé chiffrage RSA, découvert par Rivest, Shamir et Adleman en 1977.

Ève) disposant d'un ordinateur suffisamment puissant finira par casser le code, mais on peut en général se contenter de conserver secret le contenu du message pendant un temps limité. Cependant, on ne peut pas exclure que l'on dispose un jour d'algorithmes très performants pour décomposer un nombre en facteurs premiers, et de plus, si des ordinateurs quantiques (§ 11.5.2) voient le jour, les limites de la factorisation seront repoussées très loin. Heureusement, la mécanique quantique vient à point nommé pour contrecarrer les efforts des espions !

"Cryptographie quantique" est une expression médiatique, mais quelque peu trompeuse : en effet, il ne s'agit pas de chiffrer un message à l'aide de la physique quantique, mais d'utiliser celle-ci pour s'assurer que la transmission d'une clé n'a pas été espionnée. Une expression plus correcte est "distribution quantique de clé" (QKD, *Quantum Key Distribution*). La transmission d'un message, chiffré ou non, peut se faire en utilisant les deux états de polarisation linéaire orthogonaux d'un photon, par exemple $|x\rangle$ et $|y\rangle$. On peut décider d'attribuer par convention la valeur 1 à la polarisation $|x\rangle$ et la valeur 0 à la polarisation $|y\rangle$: chaque photon transporte donc un bit d'information. Tout message, chiffré ou non, peut être écrit en langage binaire, comme une suite de 0 et de 1, et le message 1001110 sera codé par Alice grâce à la séquence de photons $xyyxxxy$, qu'elle expédiera à Bob par exemple par une fibre optique. À l'aide d'un prisme polarisant, Bob sépare les photons de polarisation verticale et horizontale comme dans la figure 3.2, et deux détecteurs placés derrière le prisme lui permettent de décider si le photon était polarisé horizontalement ou verticalement : il peut donc reconstituer le message. S'il s'agissait d'un message ordinaire, il y aurait bien sûr des façons bien plus simples et efficaces de le transmettre ! Remarquons simplement que si Ève s'installe sur la fibre en connaissant la base utilisée par Alice, détecte les photons et renvoie à Bob des photons de polarisation identique à ceux expédiés par Alice, Bob ne peut pas savoir que la ligne a été espionnée. Il en serait de même pour tout dispositif fonctionnant de façon classique (c'est-à-dire sans utiliser le principe de superposition) : si l'espion prend suffisamment de précautions, il est indétectable. La raison fondamentale est que l'on peut recopier l'information classique sans la modifier, ce qui n'est pas le cas pour l'information quantique : le théorème de non-clonage quantique (exercice 3.3.4) s'y oppose.

Ce sont la mécanique quantique et le principe de superposition qui viennent au secours d'Alice et de Bob, en leur permettant de s'assurer que leur message n'a pas été intercepté. Ce message n'a pas besoin d'être long (le système de transmission par la polarisation est très peu performant). Il s'agira en général de transmettre une clé permettant de chiffrer un message ultérieur, clé qui pourra être remplacée à la demande. En effectuant un choix de base aléatoire, $\{\updownarrow, \leftrightarrow\}$ ou $\{\nwarrow\!\!\!\searrow, \nearrow\!\!\!\swarrow\}$ avec

$$|\nwarrow\!\!\!\searrow\rangle = \frac{1}{\sqrt{2}}(|\updownarrow\rangle + |\leftrightarrow\rangle) \qquad |\nearrow\!\!\!\swarrow\rangle == \frac{1}{\sqrt{2}}(|\updownarrow\rangle - |\leftrightarrow\rangle)$$

Alice envoie vers Bob quatre types de photon : polarisés suivant Ox : $\updownarrow$ et Oy : $\leftrightarrow$ comme précédemment, et polarisés suivant des axes inclinés à $\pm 45^\circ$ Ox' : $\nwarrow\!\!\!\searrow$ et Oy' : $\swarrow\!\!\!\nearrow$, correspondant respectivement aux valeurs 1 et 0 des bits. De même, Bob analyse les photons envoyés par Alice en effectuant lui aussi un choix aléatoire de base, $\{\updownarrow, \leftrightarrow\}$ ou $\{\nwarrow\!\!\!\searrow, \swarrow\!\!\!\nearrow\}$, et détermine la polarisation à l'aide d'un prisme polarisant. Une possibilité serait d'utiliser un prisme polarisant orienté aléatoirement soit verticalement, soit à 45° de la verticale et de détecter les photons sortant de ce prisme comme sur la figure 3.3. Cependant, au lieu de faire tourner l'ensemble cristal+détecteurs, on utilise plutôt un modulateur électro-optique qui permet de transformer une polarisation donnée en une polarisation orientée de façon arbitraire et de maintenir fixe l'ensemble cristal+détecteur (figure 3.4). La figure 3.5 donne un exemple d'échanges entre Alice et Bob : Bob enregistre 1 si le photon est polarisé $\updownarrow$ ou $\nwarrow\!\!\!\searrow$, 0 s'il est polarisé $\leftrightarrow$ ou $\swarrow\!\!\!\nearrow$. Après enregistrement d'un nombre suffisant de photons, Bob annonce publiquement la suite des bases qu'il a utilisées, mais non ses résultats. Alice compare sa séquence de bases à celle de Bob et lui donne toujours publiquement la liste des bases compatibles avec les siennes. Les bits qui correspondent à des analyseurs et des polariseurs incompatibles sont rejetés ($-$), et, pour les bits restants, Alice et Bob sont certains que leurs valeurs sont les mêmes : ce sont les bits qui serviront à composer la clé, et ils sont connus seulement de Bob et Alice, car l'extérieur ne connaît que la liste des orientations, pas les résultats ! Le protocole décrit ci-dessus est une réalisation possible d'un protocole appelé BB84, du nom de ses inventeurs Bennett et Brassard en 1984.

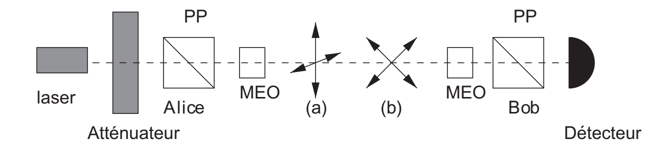


Fig. 3.4 – Schéma du protocole BB84 avec photons polarisés. Un faisceau laser est atténué de façon à simuler l'envoi de photons uniques (note 7). Un prisme polarisant PP sélectionne la polarisation, que l'on peut faire tourner à l'aide de modulateurs électro-optiques MEO. Les photons sont soit polarisés verticalement/horizontalement (a), soit à $\pm 45^\circ$ (b).

L'échange des photons polarisés constitue la partie quantique du protocole. Il reste à s'assurer que le message n'a pas été intercepté et que la clé qu'il contenait peut être utilisée sans risque. Alice et Bob choisissent au hasard un sous-ensemble de leur clé et le comparent publiquement. La conséquence de l'interception de photons par Ève serait une réduction de la corrélation entre les valeurs de leurs bits : supposons par exemple qu'Alice envoie un photon polarisé suivant Ox. Si Ève l'intercepte avec un analyseur orienté suivant Ox',

polariseurs d'Alice	↕	↔	⤢	↕	⤢	⤢	⤡	↕	⤡
séquences de bits	1	0	0	1	0	0	1	1	1
analyseurs de Bob	+	×	+	+	×	×	+	+	×
mesures de Bob	1	1	0	1	0	0	1	1	1
bits retenus	1	–	–	1	0	0	–	1	1

Fig. 3.5 – Cryptographie quantique : transmission de photons polarisés entre Bob et Alice.

et que le photon est transmis par son analyseur, elle ne sait pas que ce photon était initialement polarisé suivant Ox ; elle renvoie donc à Bob un photon polarisé dans la direction Ox', et dans 50 % des cas Bob ne va pas trouver le bon résultat. Comme Ève a une chance sur deux d'orienter son analyseur dans la bonne direction, Alice et Bob vont enregistrer une différence dans 25 % des cas et en conclure que le message a été intercepté. En résumé, la sécurité du protocole dépend du fait qu'Ève ne peut pas déterminer l'état de polarisation d'un photon si elle ne sait pas par avance dans quelle base il a été préparé.

Un fois l'échange de photons terminé, Alice et Bob disposent chacun d'une série de bits qui sont théoriquement identiques : la série 110011 dans l'exemple de la figure 3.5. Cette série de bits forme une clé, une suite aléatoire de 0 et de 1 qui peut servir pour le cryptage classique ultérieur d'un message. Elle doit bien sûr être connue uniquement d'Alice et de Bob. En pratique, il existe des sources d'erreurs provenant des imperfections des détecteurs et de la fibre optique ou des tentatives d'espionnage. Comme nous l'avons expliqué ci-dessus, Alice et Bob sacrifient un sous-ensemble commun de leur série de bits et le comparent publiquement. Ceci leur permet de mesurer le taux d'erreur par bit quantique (QBER, *Quantum Bit Error Rate*), qui est simplement la probabilité que Bob mesure une valeur erronée de la polarisation, alors qu'il connaît celle envoyée par Alice. Grâce à la connaissance de ce taux d'erreur, ils peuvent utiliser un code correcteur d'erreurs classique (par opposition à quantique) qui leur permet de reconstituer deux séries de bits strictement identiques et aléatoires, qui forment la clé secrète. Cependant, l'ensemble du processus a éventuellement permis à Ève d'acquérir une certaine information sur la série de bits. Alice et Bob doivent donc utiliser un processus également classique appelé amplification de la confidentialité (*privacy amplification*) qui leur permet, en raccourcissant leur série de bits, d'avoir la certitude qu'Ève ne possède aucune information sur la série ainsi tronquée, et c'est cette dernière qui sera utilisée en tant que clé. Dans le cas du protocole BB84, on peut montrer que le QBER doit être inférieur à 11 % si Alice et Bob veulent

disposer d'une clé fiable. Enfin, il faudrait utiliser des photons uniques, et non des paquets d'états cohérents produits par une lumière laser atténuée qui sont moins sûrs, mais que l'on choisit souvent pour des raisons pratiques[7]. Lorsque l'on utilise des fibres optiques, il est difficile de contrôler la polarisation sur de longues distances, et un support physique différent, la phase des photons, doit être utilisé pour mettre en œuvre le protocole BB84. Dans ce cas, on peut établir une clé sur environ 100 kilomètres avec un taux de l'ordre de quelques dizaines de kbits/seconde, et il existe aujourd'hui plusieurs versions commercialisées du dispositif.

La principale limitation de la cryptographie quantique vient de l'atténuation du signal dans une fibre optique. Dans le cas d'une impulsion lumineuse se propageant dans une fibre, l'intensité du signal diminue avec la distance (typiquement d'un facteur 100 sur 100 km) si bien que l'on doit utiliser des répéteurs pour retrouver la forme et l'intensité initiales du signal. Ceci n'est pas possible pour des photons uniques, qui ne peuvent pas être clonés et donc amplifiés. La distance maximale qui sépare les utilisateurs sur une ligne de cryptage quantique est aujourd'hui limitée à une centaine de km. La raison essentielle vient de l'emploi de fibres optiques et de détecteurs imparfaits. En particulier, et c'est un facteur important, les détecteurs ont des probabilités non négligeables de se déclencher alors qu'aucun photon n'est incident : c'est ce que l'on appelle les coups sombres qui limitent de façon draconienne le rapport signal/bruit de la ligne de communication. D'un point de vue technologique, la solution consiste soit à améliorer la transmission des fibres optiques, soit à diminuer la probabilité de coups sombres (dark counts) des détecteurs. On ne gagnera probablement plus grand chose sur les pertes des fibres optiques. En revanche, les détecteurs continuent à progresser mais cela relève de la physique des semiconducteurs et les problèmes sont complexes. La physique quantique apporte une solution là où on ne l'attendait pas, *via* l'utilisation du protocole de téléportation quantique (§ 11.5.3), longtemps considéré comme une curiosité de physique fondamentale. La téléportation quantique consiste à transférer à distance par exemple un état de polarisation inconnu porté par un photon sans transmettre le photon lui-même. Le succès d'une telle opération s'accompagne d'un signal électrique qui permet de déclencher les détecteurs de Bob placés en bout de ligne de communication, conditionnellement au succès du protocole. Il est alors possible de réduire le bruit apparent dans les détecteurs de Bob et donc d'augmenter le rapport signal/bruit de la ligne, au prix

---

7. Impulsions atténuées et photons uniques. Une impulsion laser atténuée utilisée en cryptographie quantique contient typiquement 0.1 photon en moyenne. On peut alors montrer qu'une impulsion non vide a une probabilité de 5 % de contenir deux photons (§ 10.2.1), un fait qui peut être exploité par Ève, mais contre lequel on peut se prémunir en utilisant un "leurre". Dans le cas de transmission de photons uniques, le théorème de non clonage quantique garantit qu'il est impossible à Ève de tromper Bob, même s'il lui est possible de faire moins de 25 % d'erreurs sur les photons renvoyés à Bob en utilisant une technique d'interception plus sophistiquée que l'interception-renvoi. Par exemple, une technique de clonage partiel permet de ramener à 16 % la probabilité de renvoyer à Bob un photon dont l'état de polarisation est incorrect.

d'un débit réduit : c'est le principe du relais quantique (voir la figure 11.19). Pour obtenir des répéteurs, il faudrait en plus utiliser des mémoires quantiques dont les réalisations pratiques semblent encore très lointaines.

3.2 Spin 1/2

3.2.1 Moment angulaire et moment magnétique en physique classique

Notre second exemple de système quantique élémentaire sera celui du spin 1/2. En l'absence d'une limite ondulatoire classique comme dans le cas du photon, la partie classique sera beaucoup plus sommaire que celle de la section précédente. Considérons une particule de masse m et de charge q décrivant une orbite fermée dans un champ de forces central (figure 3.6), $\vec{r}(t)$ et $\vec{p}(t)$ désignant la position et l'impulsion de cette particule. Soit $\mathrm{d}\vec{\mathcal{A}}$ l'aire orientée balayée par le rayon vecteur en un temps $\mathrm{d}t$, qui vérifie

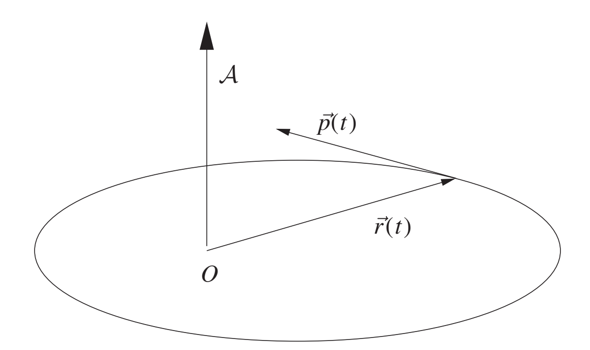


FIG. 3.6 – Facteur gyromagnétique.

$$\frac{\mathrm{d}\vec{\mathcal{A}}}{\mathrm{d}t} = \frac{1}{2m}\,\vec{r}\times\vec{p} = \frac{1}{2m}\,\vec{\jmath}$$

où $\vec{\jmath}$ est le moment angulaire. Rappelons que pour un mouvement dans un champ de forces central, ce moment angulaire est un vecteur fixé, perpendiculaire au plan de l'orbite. En intégrant sur une période, on relie l'aire totale orientée de l'orbite $\vec{\mathcal{A}}$ à $\vec{\jmath}$ et à la période T

$$\vec{\mathcal{A}} = \frac{T}{2m}\vec{\jmath}$$

Le courant induit par la charge est $I = q/T$ car la charge q passe $1/T$ fois par seconde devant un point donné, et le moment magnétique $\vec{\mu}$ induit par ce courant vaut

$$\boxed{\vec{\mu} = I\vec{\mathcal{A}} = \frac{q}{2m}\,\vec{\jmath} = \gamma\vec{\jmath}} \tag{3.30}$$

Le *facteur gyromagnétique* γ, c'est-à-dire le rapport du moment magnétique $\vec{\mu}$ au moment angulaire $\vec{\jmath}$ défini par (3.30) vaut $q/(2m)$. Le mouvement des électrons dans les atomes entraîne l'existence d'un magnétisme atomique et le mouvement des protons dans les noyaux atomiques celle d'un magnétisme nucléaire. Cependant, le mouvement des charges ne peut expliquer quantitativement ni le magnétisme atomique, ni le magnétisme nucléaire. Il faut tenir compte d'un magnétisme intrinsèque aux particules. L'expérience montre que les particules élémentaires – de spin non nul – portent un moment magnétique associé à un moment angulaire intrinsèque, appelé *spin* de la particule, que nous noterons $\vec{s}$. On peut essayer de se représenter de façon intuitive ce moment angulaire comme provenant d'une rotation de la particule sur elle-même. Cette image intuitive peut être utile, mais il ne faut pas la prendre très au sérieux : prise à la lettre, elle conduit à des contradictions insurmontables, et seule la mécanique quantique permet une description correcte du spin. L'expérience montre que l'électron, le proton et le neutron ont un spin $\frac{1}{2}\hbar$. On omet souvent le facteur $\hbar$, et on dit simplement que *l'électron, le proton et le neutron sont des particules de spin 1/2.* Le facteur gyromagnétique associé au spin est différent de (3.30). Il vaut par exemple pour l'électron[8] et le proton

$$\text{électron :}\quad \gamma_e = 2\,\frac{q_e}{2m_e} \qquad\qquad \text{proton :}\quad \gamma_p = 5.59\,\frac{q_p}{2m_p}$$

où $(q_e, q_p = -q_e)$ et (m_e, m_p) sont les charges et les masses de l'électron et du proton ; le facteur 2 pour l'électron est justifié au § 7.4.2. Mieux, bien que de charge nulle, le neutron possède un moment magnétique ! Son facteur gyromagnétique est donné par

$$\gamma_n = -3.83\,\frac{q_p}{2m_p}$$

Le magnétisme des atomes est dû à la combinaison du mouvement des électrons (magnétisme orbital) et du magnétisme associé au spin des électrons. Le magnétisme des noyaux atomiques est dû au mouvement des protons et au magnétisme associé aux spins des neutrons et des protons. L'équation (3.30) montre que le facteur gyromagnétique est inversement proportionnel à la masse : le magnétisme d'origine nucléaire est plus faible que le magnétisme d'origine électronique par un facteur $\sim m_e/m_p \sim 1/1000$. Malgré ce facteur défavorable, le magnétisme nucléaire joue un rôle pratique considérable en étant à la base de la résonance magnétique nucléaire (RMN : section 5.2) et de ses dérivés comme l'imagerie par résonance magnétique (IRM).

Examinons en physique classique le mouvement d'un moment magnétique $\vec{\mu}$ dans un champ magnétique constant $\vec{B}$. Ce moment magnétique est soumis

8. À des corrections près de l'ordre de 0.1 % : ces corrections sont calculables grâce à l'électrodynamique quantique.

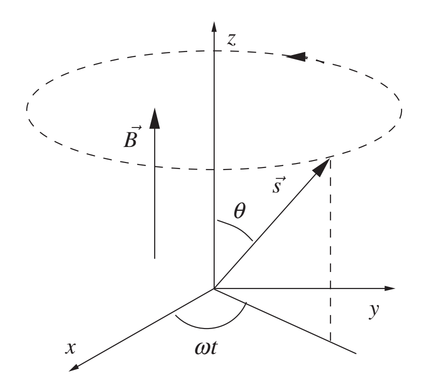


FIG. 3.7 – Précession de Larmor : le spin $\vec{s}$ précesse autour de $\vec{B}$ avec une fréquence angulaire ω.

à un couple $\vec{\Gamma} = \vec{\mu} \times \vec{B}$, et l'équation du mouvement est

$$\frac{\mathrm{d}\vec{s}}{\mathrm{d}t} = \vec{\mu} \times \vec{B} = \frac{q}{2m}\,\vec{s} \times \vec{B} = -\frac{qB}{2m}\,\hat{B} \times \vec{s} \tag{3.31}$$

Cette équation implique que $\vec{s}$ et $\vec{\mu}$ tournent autour de $\vec{B}$ avec une vitesse angulaire constante $\omega = -qB/(2m)$, appelée *fréquence de Larmor*. Il est commode de donner une valeur algébrique à ω : la rotation se fait dans le sens trigonométrique pour $q < 0$ ($\omega > 0$). Le mouvement de rotation est appelé précession de Larmor (figure 3.7).

3.2.2 Expérience de Stern-Gerlach et filtres de Stern-Gerlach

L'expérience réalisée par Stern et Gerlach en 1921 est schématisée sur la figure 3.8. Un jet d'atomes d'argent sort d'un four et est collimaté par deux fentes, avant de passer dans l'entrefer d'un aimant où règne un champ magnétique dirigé suivant[9] Oz. Le champ magnétique est inhomogène : B_z est une fonction de z. L'atome d'argent porte un moment magnétique qui est en fait le moment magnétique de son électron de valence. Du point de vue des forces magnétiques, tout se passe comme si un électron traversait l'entrefer de l'aimant. Cependant, on doit utiliser dans la dynamique la masse de l'atome et non celle de l'électron et noter l'absence de force de Lorentz, l'atome d'argent étant électriquement neutre. L'énergie potentielle U d'un moment magnétique dans $\vec{B}$ est $U = -\vec{\mu} \cdot \vec{B}$, et la force correspondante

9. Le lecteur prendra garde au fait que l'orientation des axes est différente de celle de la section précédente : la direction de propagation est maintenant Oy. Ce nouveau choix est dicté par le souhait de respecter les conventions usuelles.

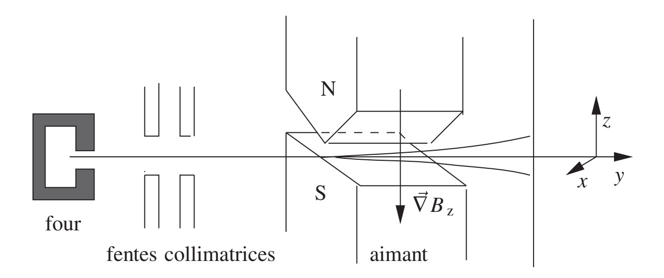


FIG. 3.8 – Expérience de Stern-Gerlach.

$$\vec{F} = -\vec{\nabla} U \qquad F_z = \mu_z \frac{\partial B_z}{\partial z} \tag{3.32}$$

En réalité, $\vec{B}$ ne peut pas être strictement parallèle à Oz : si $\vec{B} = (0, 0, B)$, $\partial B/\partial z \neq 0$ est incompatible avec l'équation de Maxwell $\vec{\nabla} \cdot \vec{B} = 0$. Une justification complète de (3.32) se trouve dans l'exercice 8.6.12, où l'on montre que la force effective sur l'atome est bien donnée par (3.32). Lorsque le champ magnétique est nul, les atomes arrivent au voisinage d'un point de l'écran et forment une tache de dimension finie en raison de la dispersion des vitesses, car la collimation n'est pas parfaite. L'orientation des moments magnétiques à la sortie du four est *a priori* aléatoire, et en présence du champ magnétique, on s'attend à un élargissement de la tache : les atomes dont le moment magnétique $\vec{\mu}$ est antiparallèle à Oz subissent une déviation maximale vers le haut pour $(\partial B_z/\partial z) < 0$, ceux dont $\vec{\mu}$ est parallèle à Oz une déviation maximale vers le bas, toutes les déviations intermédiaires étant possibles. Ce résultat est contredit par l'expérience : on observe *deux* taches symétriques par rapport au point d'arrivée en l'absence de champ magnétique. Tout se passe comme si μ_z, et donc s_z, ne pouvait prendre que deux valeurs, et deux seulement, dont on constate[10] qu'elles correspondent à $s_z = \pm\hbar/2$: s_z est *quantifié*. On remarquera que comme le facteur gyromagnétique est négatif ($\gamma < 0$), la déviation vers le haut (resp. bas) correspond à $s_z > 0$ (resp. < 0). L'appareil de Stern-Gerlach agit comme la lame biréfringente de la figure 3.2 : à la sortie de l'appareil, l'électron suit une trajectoire[11] où son spin est orienté soit vers le haut : $s_z = +\hbar/2$, soit vers le bas : $s_z = -\hbar/2$. L'analogie avec la polarisation des photons nous suggère de prendre comme espace des états de spin $1/2$ un espace vectoriel à deux dimensions, ce qui s'avèrera être le bon choix. Une base possible de cet espace est formée des deux vecteurs $|+\rangle$ et $|-\rangle$, décrivant les états physiques obtenus en sélectionnant les atomes déviés vers le haut ou vers le bas par l'appareil de Stern-Gerlach, et correspondant respectivement aux valeurs $+\hbar/2$ et $-\hbar/2$ de s_z. Les états $|+\rangle$ et $|-\rangle$ sont souvent appelés

10. La connaissance de $\partial B_z/\partial z$ et de γ permet en principe de remonter à la valeur de s_z à partir de la déviation : exercice 8.6.12.

11. On peut montrer (exercice 8.6.12) que les trajectoires peuvent être traitées classiquement.

“spin up” et “spin down”. Ces états de spin sont l’analogue de deux états de polarisation orthogonale $|\Phi\rangle$ et $|\Phi_\perp\rangle$ dans le cas des photons[12].

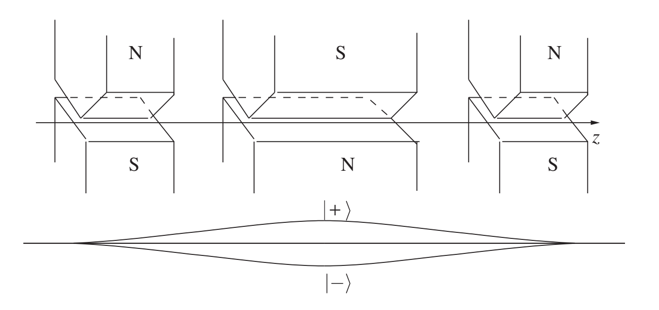


FIG. 3.9 – Filtre de Stern-Gerlach.

Le dispositif schématisé sur la figure 3.9 permet de recombiner les atomes déviés vers le haut ou vers le bas sur une trajectoire unique, de même que la combinaison de deux lames biréfringentes de la figure 3.3 permettait de recombiner les trajectoires des photons polarisés suivant Ox et suivant Oy. Ce dispositif, que nous appellerons “filtre de Stern-Gerlach” n’a pas été réalisé expérimentalement par Stern et Gerlach. Il a été imaginé 40 ans plus tard par Wigner pour les besoins d’une discussion théorique. Si l’on place deux filtres de Stern-Gerlach à la suite l’un de l’autre avec la même orientation de $\vec{B}$ en bloquant par exemple les deux voies du bas (figure 3.10a), on constate que 100 % des atomes qui passent le premier filtre sont aussi transmis par le second, de même qu’un photon sélectionné par un polariseur orienté suivant Ox est transmis avec une probabilité de 100 % par un analyseur de même orientation. Si au contraire la voie du bas est bloquée sur le premier filtre et la voie du haut sur le second (figure 3.10b), alors aucun atome n’est transmis, de même qu’aucun photon n’est transmis si l’analyseur et le polariseur sont croisés. Comme dans la section précédente, on rend compte de ces résultats en écrivant les amplitudes de probabilité $a(+ \to +)$ et $a(+ \to -)$ comme des produits scalaires de vecteurs de base[13]

$$a(+ \to +) = \langle +|+\rangle = 1 \quad a(- \to -) = \langle -|-\rangle = 1 \quad a(+ \to -) = \langle -|+\rangle = 0 \tag{3.33}$$

Si l’on représente les vecteurs $|+\rangle$ et $|-\rangle$ sous la forme de vecteurs colonnes

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.34}$$

12. Toutefois, il ne faut pas pousser trop loin cette analogie ; comme nous le verrons au chapitre 9, le photon a un spin $\hbar$, et non $\hbar/2$. Un spin $\hbar$ a normalement trois états de polarisation possibles. Il y en a seulement deux dans le cas du photon parce que le photon a une masse nulle.

13. En toute rigueur, on sait seulement que $|a(+ \to +)| = |a(- \to -)| = 1$, mais un choix de phase convenable permet toujours de se ramener à (3.33).

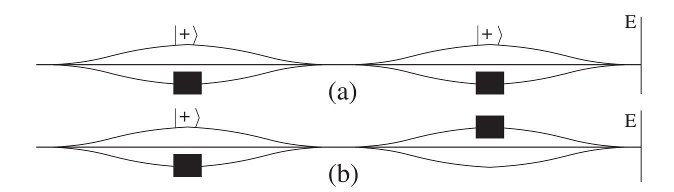


FIG. 3.10 – Filtres de Stern-Gerlach en série. En (b), aucun atome n'arrive sur l'écran.

le vecteur d'état (unitaire) le plus général $|\chi\rangle \in \mathcal{H}$ s'écrira

$$|\chi\rangle = \lambda|+\rangle + \mu|-\rangle \quad \text{ou} \quad |\chi\rangle = \begin{pmatrix} \lambda \\ \mu \end{pmatrix} \qquad |\lambda|^2 + |\mu|^2 = 1 \tag{3.35}$$

Avec les vecteurs $|+\rangle$ et $|-\rangle$ on peut construire un opérateur hermitien S_z tel que ces vecteurs soient vecteurs propres de S_z avec les valeurs propres $\pm\hbar/2$

$$S_z = \frac{1}{2}\,\hbar\Big(|+\rangle\langle+| - |-\rangle\langle-|\Big) = \frac{1}{2}\,\hbar\Big(\mathcal{P}_+ - \mathcal{P}_-\Big) = \frac{1}{2}\,\hbar\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.36}$$

où $\mathcal{P}_+$ et $\mathcal{P}_-$ sont les projecteurs sur les états $|+\rangle$ et $|-\rangle$. À la propriété physique $\mathcal{S}_z$, composante suivant z du spin, on associe un opérateur hermitien S_z agissant dans l'espace des états $\mathcal{H}$. Les vecteurs $|+\rangle$ et $|-\rangle$ sont aussi appelés *états propres de* S_z, et forment *la base où* S_z *est diagonal* : dans cette base S_z est représenté par la matrice diagonale (3.36). La propriété physique : composante suivant z du spin, a une valeur bien déterminée $+\hbar/2$ ou $-\hbar/2$ si le vecteur d'état $|\chi\rangle$ est égal à $|+\rangle$ ou $|-\rangle$.

3.2.3 États de spin d'orientation arbitraire

Poursuivons l'analogie avec la polarisation d'un photon en faisant tourner la direction du champ magnétique du filtre de Stern-Gerlach et en l'alignant dans la direction $\hat{n}$: seule la composante $B_{\hat{n}} = \vec{B} \cdot \hat{n}$ du champ magnétique est non nulle. Avec cette nouvelle orientation, le filtre de Stern-Gerlach va fabriquer des états que nous noterons $|+, \hat{n}\rangle$ et $|-, \hat{n}\rangle$, obtenus en sélectionnant les atomes déviés respectivement dans le sens de $\hat{n}$ et dans la direction opposée[14]. Par analogie avec le cas des photons, nous dirons que le spin 1/2 est *polarisé* dans la direction $+\hat{n}$ ou $-\hat{n}$. Nous procédons comme pour l'étude de la polarisation d'un photon, en utilisant un premier filtre de Stern-Gerlach, jouant le rôle de polariseur, dont le champ magnétique orienté suivant Oz sélectionne les spins dans l'état $|+\rangle$. Le deuxième filtre a son champ magnétique orienté dans la direction $\hat{n}$ et joue le rôle d'analyseur. Il permet de

14. Ceci suppose que l'on sache changer la direction de propagation des atomes d'argent pour la rendre orthogonale à $\hat{n}$. Comme nous discutons une "expérience théorique", nous ne nous attarderons pas sur les moyens qui pourrraient être utilisés pour ce faire.

mesurer expérimentalement les probabilités $\mathsf{p}(+ \to [+,\hat{n}]) = |\langle +,\hat{n}|+\rangle|^2$ et $\mathsf{p}(+ \to [-,\hat{n}]) = |\langle -,\hat{n}|+\rangle|^2$; comme dans la section précédente, nous supposons que ces probabilités sont données par le module carré de produits scalaires. De même que les états[15] $|+\rangle$ et $|-\rangle$, les états $|+,\hat{n}\rangle$ et $|-,\hat{n}\rangle$ sont orthogonaux : $\langle +,\hat{n}|-,\hat{n}\rangle = 0$. Si le polariseur et l'analyseur sont orientés dans la même direction, un état préparé par le polariseur est transmis à 100 % par l'analyseur, et à 0 % si leurs orientations sont opposées[16] : le résultat du test de la polarisation est certain. Si les directions ne sont pas les mêmes, on observe seulement une certaine probabilité de transmission. De même que les bases d'états de polarisation d'un photon $\{|x\rangle, |y\rangle\}$ et $\{|\theta\rangle, |\theta_\perp\rangle\}$ étaient incompatibles (§ 3.1.2), les bases $\{|+\rangle, |-\rangle\}$ et $\{|+,\hat{n}\rangle, |-,\hat{n}\rangle\}$ sont incompatibles pour les états de spin 1/2 si $\hat{n} \neq \hat{z}$.

Nous allons maintenant utiliser l'invariance par rotation pour déterminer les probabilités de transmission : la physique de ce problème ne doit pas dépendre de l'orientation du système d'axes. La première conséquence de cette invariance est que la direction Oz n'a aucune raison d'être privilégiée, et qu'il doit exister un opérateur hermitien $S_{\hat{n}} = \vec{S}\cdot\hat{n}$, projection du spin sur l'axe $\hat{n}$, ayant des valeurs propres $\hbar/2$ et $-\hbar/2$ et la forme (3.36) dans une base $\{|+,\hat{n}\rangle, |-,\hat{n}\rangle\}$ qui reste à déterminer. L'opérateur $S_{\hat{n}}$ s'écrira en fonction de ses valeurs propres et de ses vecteurs propres

$$S_{\hat{n}} = \frac{1}{2}\,\hbar\Big(|+,\hat{n}\rangle\langle +,\hat{n}| - |-,\hat{n}\rangle\langle -,\hat{n}|\Big) \tag{3.37}$$

Introduisons la notion de *valeur moyenne* de la composante suivant $\hat{n}$ du spin, que nous noterons $\langle S_{\hat{n}}\rangle$. Comme la déviation dans la direction $\pm\hat{n}$ correspond à une valeur $s_{\hat{n}} = \pm\hbar/2$ lorsque le spin est dans un état $|\chi\rangle$ arbitraire, cette valeur moyenne, notée $\langle S_{\hat{n}}\rangle$, sera donnée par

$$\begin{aligned}
\langle S_{\hat{n}}\rangle &= \frac{1}{2}\,\hbar\Big(\mathsf{p}(\chi \to [+,\hat{n}]) - \mathsf{p}(\chi \to [-,\hat{n}]\Big) \\
&= \frac{1}{2}\,\hbar\Big(\langle\chi|+,\hat{n}\rangle\langle +,\hat{n}|\chi\rangle - \langle\chi|-,\hat{n}\rangle\langle -,\hat{n}|\chi\rangle\Big) \\
&= \langle\chi|\frac{1}{2}\,\hbar\Big(|+,\hat{n}\rangle\langle +,\hat{n}| - |-,\hat{n}\rangle\langle -,\hat{n}|\Big)|\chi\rangle \\
&= \langle\chi|S_{\hat{n}}|\chi\rangle
\end{aligned} \tag{3.38}$$

La matrice représentative de $S_{\hat{n}}$ dans la base (3.34) où S_z est diagonal est *a priori* donnée par la matrice hermitienne 2×2 la plus générale de valeurs propres $\pm\hbar/2$

$$S_{\hat{n}} = \frac{1}{2}\,\hbar\begin{pmatrix} a & b \\ b^* & c \end{pmatrix} = \frac{1}{2}\,\hbar A \tag{3.39}$$

où a et c sont des nombres réels. L'équation aux valeurs propres $\lambda_\pm$ de la matrice A s'écrit

$$\lambda^2 - (a+c)\lambda + ac - |b|^2 = 0$$

15. Les notations $|+\rangle$ et $|-\rangle$ sont donc des notations abrégées pour $|+,\hat{z}\rangle$ et $|-,\hat{z}\rangle$.
16. Et non orthogonales comme dans le cas des photons !

On doit avoir $\lambda_+ + \lambda_- = 0$ et $\lambda_+\lambda_- = -1$ soit

$$a + c = 0 \quad ac - |b|^2 = -1 \Rightarrow a^2 + |b|^2 = 1$$

Paramétrons a et b à l'aide de deux angles α et β : $a = \cos\beta$ et $b = \exp(-\mathrm{i}\alpha)\sin\beta$. Nous obtenons pour $S_{\hat{n}}$

$$S_{\hat{n}} = \frac{1}{2}\hbar \begin{pmatrix} \cos\beta & \mathrm{e}^{-\mathrm{i}\alpha}\sin\beta \\ \mathrm{e}^{\mathrm{i}\alpha}\sin\beta & -\cos\beta \end{pmatrix} \tag{3.40}$$

dont les vecteurs propres sont à un facteur de phase près (*cf.* (2.35))

$$|+,\hat{n}\rangle = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\alpha/2}\cos\beta/2 \\ \mathrm{e}^{\mathrm{i}\alpha/2}\sin\beta/2 \end{pmatrix} \qquad |-,\hat{n}\rangle = \begin{pmatrix} -\mathrm{e}^{-\mathrm{i}\alpha/2}\sin\beta/2 \\ \mathrm{e}^{\mathrm{i}\alpha/2}\cos\beta/2 \end{pmatrix} \tag{3.41}$$

3.2.4 Rotation d'un spin 1/2

Il nous reste à trouver une interprétation géométrique aux angles α et β. Nous allons faire l'hypothèse que la valeur moyenne $\langle\vec{S}\rangle$, dont les composantes sont $(\langle S_x\rangle, \langle S_y\rangle, \langle S_z\rangle)$, se transforme par rotation comme un vecteur de l'espace à trois dimensions, c'est-à-dire comme l'objet classique $\vec{s}$ correspondant. Reprenons l'expérience type polariseur/analyseur. Dans un premier temps, le champ magnétique du polariseur est orienté suivant Oz, et de même pour l'analyseur. Nous savons que dans ce cas 100 % des spins traversent l'analyseur. Si le champ de l'analyseur est orienté antiparallèlement à Oz, alors aucun spin ne le traverse. Nous pouvons exprimer ce résultat sous la forme suivante : à la sortie du polariseur, la valeur moyenne de S_z, $\langle S_z\rangle$, est égale à $\hbar/2$. Orientons maintenant le champ magnétique de l'analyseur suivant Ox : on constate expérimentalement que les spins ont alors une chance sur deux d'être déviés vers les x positifs et une chance sur deux d'être déviés vers les x négatifs, ce qui correspond à une valeur moyenne nulle de S_x : $\langle S_x\rangle = 0$. Ce résultat ne doit pas surprendre. Un premier argument fait appel à un raisonnement classique : un spin classique parallèle à Oz n'est pas dévié par un gradient de champ suivant Ox. Un deuxième argument plus général fait appel à l'invariance par rotation[17] : dans notre problème, les variables de spin sont découplées des variables spatiales liées à la propagation de l'atome et, pour les rotations du spin, le problème est invariant par rotation autour de Oz : en l'absence de direction privilégiée dans le plan xOy, $\langle S_x\rangle = \langle S_y\rangle = 0$. Le vecteur $\langle\vec{S}\rangle$ a donc pour composantes $(0, 0, \hbar/2)$.

Supposons maintenant que l'expérimentateur décide d'utiliser un système d'axes $x'Oz'$ obtenu à partir de xOz par une rotation d'angle $-\theta$ autour de Oy (figure 3.11a). Si $\langle\vec{S}\rangle$ est un vecteur, ses composantes dans le nouveau système d'axes seront $\hbar/2(\sin\theta, 0, \cos\theta)$. On obtient une situation physique

17. On peut aussi invoquer l'invariance par parité sans faire appel au découplage des variables de spin et des variables spatiales : voir exercice 8.6.12.

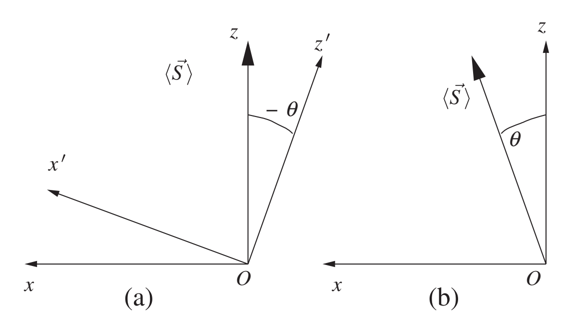


FIG. 3.11 – (a) $\langle\vec{S}\rangle$ dans deux systèmes d'axes. (b) Rotation de $\langle\vec{S}\rangle$.

équivalente[18] en conservant le système d'axes original et en orientant le gradient du champ magnétique du polariseur suivant une direction faisant un angle θ avec Oz (figure 3.11b). Le polariseur prépare alors les spins dans un état que nous noterons $|+,\hat{n}_\theta\rangle$. On a donc pour les valeurs moyennes

$$\langle S_x\rangle = \langle +,\hat{n}_\theta|S_x|+,\hat{n}_\theta\rangle = \frac{\hbar}{2}\sin\theta \qquad \langle S_z\rangle = \langle +,\hat{n}_\theta|S_z|+,\hat{n}_\theta\rangle = \frac{\hbar}{2}\cos\theta \tag{3.42}$$

En général, on pourra orienter le champ magnétique $\vec{B}$ du polariseur suivant une direction quelconque $\hat{n}$: le polariseur prépare les spins dans l'état $|+,\hat{n}\rangle$. Soit θ et ϕ les angles polaire et azimutal définissant la direction de $\hat{n}$ (figure 3.12). La généralisation immédiate de l'argument précédent montre que les valeurs moyennes de $\vec{S}$ sont alors

$$\begin{aligned}
\langle S_x\rangle &= \langle +,\hat{n}|S_x|+,\hat{n}\rangle = \frac{\hbar}{2}\sin\theta\cos\phi = \frac{\hbar}{2}\,n_x \\
\langle S_y\rangle &= \langle +,\hat{n}|S_y|+,\hat{n}\rangle = \frac{\hbar}{2}\sin\theta\sin\phi = \frac{\hbar}{2}\,n_y \\
\langle S_z\rangle &= \langle +,\hat{n}|S_z|+,\hat{n}\rangle = \frac{\hbar}{2}\cos\theta = \frac{\hbar}{2}\,n_z
\end{aligned} \tag{3.43}$$

ou bien, en notation vectorielle

$$\langle\vec{S}\rangle = \langle +,\hat{n}|\vec{S}|+,\hat{n}\rangle = \frac{\hbar}{2}\,\hat{n} \tag{3.44}$$

Nous avons détaillé le raisonnement menant à (3.44), mais nous aurions pu arriver directement au résultat en remarquant que le seul vecteur à notre disposition est $\hat{n}$, et $\langle\vec{S}\rangle$ est nécessairement parallèle à $\hat{n}$. Calculons maintenant les valeurs moyennes compte tenu de (3.41)

$$\langle S_z\rangle = \frac{\hbar}{2}\left(\cos^2\beta/2 - \sin^2\beta/2\right) = \frac{\hbar}{2}\cos\beta$$

18. Nous verrons au § 7.1.1 que ceci consiste à passer du point de vue passif au point de vue actif pour une opération de symétrie.

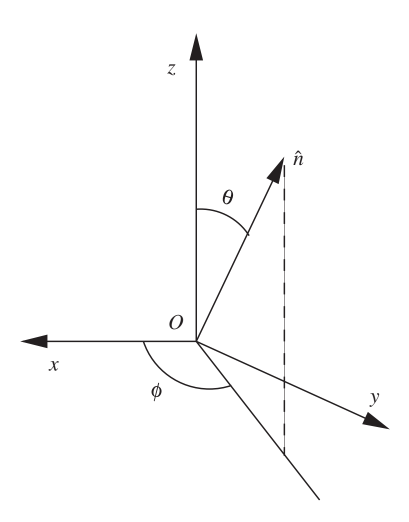


Fig. 3.12 – Orientation de $\hat{n}$.

On doit donc avoir $\beta = \pm\theta$. Choisissons la solution $\beta = \theta$ et calculons les matrices représentatives de S_x et S_y dans la base (3.34) ; comme $\theta = \beta = \pi/2$ dans les deux cas, (3.40) devient

$$S_x = \frac{1}{2}\,\hbar \begin{pmatrix} 0 & \mathrm{e}^{-\mathrm{i}\alpha_x} \\ \mathrm{e}^{\mathrm{i}\alpha_x} & 0 \end{pmatrix} \qquad S_y = \frac{1}{2}\,\hbar \begin{pmatrix} 0 & \mathrm{e}^{-\mathrm{i}\alpha_y} \\ \mathrm{e}^{\mathrm{i}\alpha_y} & 0 \end{pmatrix}$$

Ceci donne pour les valeurs moyennes

$$\langle S_x \rangle = \frac{1}{2}\,\hbar \sin\theta \cos(\alpha - \alpha_x) \qquad \langle S_y \rangle = \frac{1}{2}\,\hbar \sin\theta \cos(\alpha - \alpha_y)$$

On obtient par identification avec (3.43)

$$\cos(\alpha - \alpha_x) = \cos\phi \qquad \cos(\alpha - \alpha_y) = \sin\phi \tag{3.45}$$

La solution de (3.45) n'est pas unique[19] ; nous choisirons par convention

$$\alpha_x = 0 \qquad \alpha_y = \pi/2$$

Avec ce choix, $\alpha = \phi$ et les opérateurs S_x, S_y et S_z dans la base (3.34) prennent la forme

$$S_x = \frac{1}{2}\,\hbar\sigma_x \qquad S_y = \frac{1}{2}\,\hbar\sigma_y \qquad S_z = \frac{1}{2}\,\hbar\sigma_z \tag{3.46}$$

Les matrices σ_x, σ_y et σ_z sont appelées *matrices de Pauli*

$$\boxed{\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \tag{3.47}$$

19. Les autres solutions correspondent à un système d'axes obtenu par rotation autour de Oz des axes Ox et Oy ou à un système d'axes obtenu par inversion de Oy : *cf.* exercice 3.3.6.

Ces matrices vérifient des relations importantes souvent utilisées

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I \qquad \sigma_x\sigma_y = \mathrm{i}\sigma_z \quad \text{et permutations} \tag{3.48}$$

que l'on peut rassembler en

$$\boxed{\sigma_i\sigma_j = \delta_{ij} + \mathrm{i}\sum_k \varepsilon_{ijk}\sigma_k} \tag{3.49}$$

où les indices (i, j, k) prennent les valeurs (x, y, z) et ε_{ijk} est le *tenseur complètement antisymétrique*, égal à $+1$ si (ijk) est une permutation paire de (xyz), à -1 dans le cas d'une permutation impaire et à zéro dans tous les autres cas[20]. Une forme équivalente de (3.49) est la suivante : si $\vec{a}$ et $\vec{b}$ sont deux vecteurs

$$(\vec{\sigma}\cdot\vec{a})(\vec{\sigma}\cdot\vec{b}) = \vec{a}\cdot\vec{b} + \mathrm{i}\vec{\sigma}\cdot(\vec{a}\times\vec{b}) \tag{3.50}$$

où l'on a utilisé pour le produit vectoriel

$$(\vec{a}\times\vec{b})_i = \sum_{jk} \varepsilon_{ijk} a_j b_k \tag{3.51}$$

L'équation (3.49) implique aussi les relations de commutation[21]

$$[\sigma_i, \sigma_j] = 2\mathrm{i}\sum_k \varepsilon_{ijk}\sigma_k \tag{3.52}$$

ou, de façon équivalente pour les composantes du spin

$$\boxed{[S_i, S_j] = \mathrm{i}\,\hbar\sum_k \varepsilon_{ijk} S_k} \tag{3.53}$$

Les matrices de Pauli forment avec la matrice identité I une base pour l'espace vectoriel des matrices sur $\mathcal{H}$. En effet toute matrice 2×2 peut s'écrire

$$A = \lambda_0 I + \sum_i \lambda_i\sigma_i \tag{3.54}$$

où les coefficients λ_0 et λ_i sont réels pour une matrice hermitienne $A = A^\dagger$; ils sont donnés par (exercice 3.3.7)

$$\lambda_0 = \frac{1}{2}\,\mathrm{Tr}\,A \qquad \lambda_i = \frac{1}{2}\,\mathrm{Tr}\,A\sigma_i \tag{3.55}$$

Le fait que les matrices de Pauli forment une base pour les matrices sur tout espace de Hilbert à deux dimensions entraîne que ces matrices sont souvent

20. Par exemple $\varepsilon_{yzx} = 1$, $\varepsilon_{yxz} = -1$ et $\varepsilon_{xxz} = 0$.

21. En écrivant les indices explicitement : $[\sigma_x, \sigma_y] = 2\mathrm{i}\,\sigma_z$ et deux autres relations obtenues par permutation circulaire des indices (x, y, z).

utilisées pour des problèmes où l'espace des états est à deux dimensions, même si le problème physique n'a rien à voir avec un spin 1/2. Elles sont par exemple très utiles pour traiter un modèle standard de la physique atomique, l'atome dit "à deux niveaux" (voir la section 5.4 et le § 15.3.1).

Les vecteurs propres $|+,\hat{n}\rangle$ et $|-,\hat{n}\rangle$ de $S_{\hat{n}} = \frac{1}{2}\,\hbar\vec{\sigma}\cdot\hat{n}$ se déduisent de (3.41) avec $\beta = \theta$ et $\alpha = \varphi$

$$|+,\hat{n}\rangle = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\phi/2}\cos\theta/2 \\ \mathrm{e}^{\mathrm{i}\phi/2}\sin\theta/2 \end{pmatrix} \qquad |-,\hat{n}\rangle = \begin{pmatrix} -\mathrm{e}^{-\mathrm{i}\phi/2}\sin\theta/2 \\ \mathrm{e}^{\mathrm{i}\phi/2}\cos\theta/2 \end{pmatrix} \tag{3.56}$$

Les états $|+,\hat{n}\rangle$ et $|-,\hat{n}\rangle$ sont les transformés des états $|+\rangle$ et $|-\rangle$ par une rotation qui amène l'axe Oz sur l'axe $\hat{n}$: un choix possible, cohérent avec celui qui sera fait ultérieurement au chapitre 9, consiste à effectuer une première rotation de θ autour de Oy, suivie d'une rotation de ϕ autour de Oz. On peut écrire (3.56) sous la forme

$$\begin{aligned} |+,\hat{n}\rangle &= D^{(1/2)}_{++}(\theta,\phi)|+\rangle + D^{(1/2)}_{-+}(\theta,\phi)|-\rangle \\ |-,\hat{n}\rangle &= D^{(1/2)}_{+-}(\theta,\phi)|+\rangle + D^{(1/2)}_{--}(\theta,\phi)|-\rangle \end{aligned} \tag{3.57}$$

Cette équation définit une matrice[22] $D^{(1/2)}(\theta,\phi)$, appelée *matrice de rotation* pour le spin 1/2

$$D^{(1/2)}(\theta,\phi) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\phi/2}\,\cos\theta/2 & -\mathrm{e}^{-\mathrm{i}\phi/2}\,\sin\theta/2 \\ \mathrm{e}^{\mathrm{i}\phi/2}\,\sin\theta/2 & \mathrm{e}^{\mathrm{i}\phi/2}\,\cos\theta/2 \end{pmatrix} \tag{3.58}$$

Cette matrice est unitaire, car elle effectue un changement de base dans $\mathcal{H}$, et de plus on vérifie qu'elle est de déterminant 1, et c'est donc une matrice appartenant au groupe $SU(2)$: *cf.* exercice 7.5.2. Il est intéressant de considérer les rotations de 2π, qui ramènent le système physique à sa position initiale. On remarque que par exemple $D^{1/2}(\theta = 2\pi, \phi = 0) = -I$. Dans une rotation de 2π autour de Oy, le vecteur d'état $|\chi\rangle \to -|\chi\rangle$! Mais il n'y a là aucun paradoxe : les vecteurs $|\chi\rangle$ et $-|\chi\rangle$ représentent le même état physique, et, comme il se doit, une rotation de 2π ne modifie pas l'état physique. Ce comportement du spin 1/2 est à contraster avec celui des photons : d'après (3.28), $\exp(-2\mathrm{i}\pi\Sigma_z) = +I$, et le vecteur d'état est inchangé dans une rotation de 2π. Nous observons là une différence remarquable entre spins entiers et spins demi-entiers, sur laquelle nous reviendrons au chapitre 9.

Nous allons mettre la matrice de rotation $D^{(1/2)}(\theta,\phi)$ sous une forme qui nous sera très utile pour la suite. Appelons $\mathcal{R}_{\hat{p}}(\theta)$ la rotation dans $\mathbb{R}^3$ d'angle θ autour d'un vecteur unitaire $\hat{p}$. Nous allons montrer que l'opérateur unitaire de rotation $U[\mathcal{R}_{\hat{p}}(\theta)]$ pour un spin 1/2, est donné par

$$\boxed{U[\mathcal{R}_{\hat{p}}(\theta)] = \exp\left(-\mathrm{i}\frac{\theta}{2}\,\vec{\sigma}\cdot\hat{p}\right) = I\,\cos\frac{\theta}{2} - \mathrm{i}(\vec{\sigma}\cdot\hat{p})\sin\frac{\theta}{2}} \tag{3.59}$$

22. On remarque que cette matrice s'écrit en fonction de $\theta/2$, et non de θ comme dans le cas d'un photon (3.28) : le photon a un spin 1 et non 1/2 !

Pour établir la seconde identité, on remarque que $(\vec{\sigma}\cdot\hat{p})^2 = 1$ d'après (3.50), et en développant l'exponentielle

$$\exp\left(-\mathrm{i}\frac{\theta}{2}\,\vec{\sigma}\cdot\hat{p}\right) = I + \frac{-\mathrm{i}\theta}{2}\vec{\sigma}\cdot\hat{p} + \frac{1}{2!}\left(\frac{-\mathrm{i}\theta}{2}\right)^2 I + \frac{1}{3!}\left(\frac{-\mathrm{i}\theta}{2}\right)^3 \vec{\sigma}\cdot\hat{p} + \cdots$$

On reconnaît en facteur de I le développement de $\cos\theta/2$ et en facteur de $-\mathrm{i}\vec{\sigma}\cdot\hat{p}$ celui de $\sin\theta/2$, ce qui montre la seconde égalité dans (3.59). Pour vérifier que l'opérateur $U[\mathcal{R}_{\hat{p}}(\theta)]$ est bien l'opérateur qui effectue une rotation de θ autour de l'axe $\hat{p}$, prenons $\hat{p} = (-\sin\phi, \cos\phi, 0)$: une rotation de θ autour de cet axe amène l'axe Oz sur $\hat{n}$. Avec ce choix de $\hat{p}$, nous écrivons sous forme matricielle

$$\exp\left(-\mathrm{i}\frac{\theta}{2}\,\vec{\sigma}\cdot\hat{p}\right) = \begin{pmatrix} \cos\theta/2 & -\mathrm{e}^{-\mathrm{i}\phi}\,\sin\theta/2 \\ \mathrm{e}^{\mathrm{i}\phi}\,\sin\theta/2 & \cos\theta/2 \end{pmatrix} \tag{3.60}$$

Cette matrice ne semble pas coïncider avec la matrice (3.58), mais la différence est absorbée dans des facteurs de phase qui ne sont pas physiquement pertinents. On peut aisément vérifier que

$$U[\mathcal{R}_{\hat{p}}(\theta)]|+\rangle = \mathrm{e}^{\,\mathrm{i}\phi/2}|+,\hat{n}\rangle \qquad U[\mathcal{R}_{\hat{p}}(\theta)]|-\rangle = \mathrm{e}^{-\mathrm{i}\phi/2}|-,\hat{n}\rangle$$

ou remarquer que

$$U[\mathcal{R}_{\hat{p}}(\theta)] = D^{(1/2)}(\theta,\phi)\,\mathrm{e}^{\,\mathrm{i}\phi\sigma_z/2} = D^{(1/2)}(\theta,\phi)U[\mathcal{R}_{\hat{z}}(-\phi)]$$

et qu'une rotation de $-\phi$ autour de Oz laisse invariant l'axe Oz : dans cette opération, les états $|+\rangle$ et $|-\rangle$ sont multipliés par un facteur de phase. Il est commode de visualiser géométriquement un état de spin $|\chi\rangle$ en se donnant un vecteur unitaire, le *vecteur de Bloch* $\vec{b}$, dont on trouvera la généralisation au § 11.1.3, qui définit un point sur la sphère de rayon unité, la *sphère de Poincaré-Bloch*

$$\vec{b} = \langle\vec{\sigma}\rangle = \langle\chi|\vec{\sigma}|\chi\rangle$$

Ce vecteur est représenté sur la figure 3.13 dans les cas suivants

$$|\chi\rangle = |\pm\rangle \qquad |\chi\rangle = |+,\hat{x}\rangle \qquad |\chi\rangle = |+,\hat{n}\rangle$$

La forme (3.56) des vecteurs propres de $S_{\hat{n}}$ permet de calculer les amplitudes de probabilité

$$\begin{aligned} a(+\to[+,\hat{n}]) &= \langle +,\hat{n}|+\rangle = \cos\theta/2\,\mathrm{e}^{\mathrm{i}\phi/2} \\ a(+\to[-,\hat{n}]) &= \langle -,\hat{n}|+\rangle = -\sin\theta/2\,\mathrm{e}^{\mathrm{i}\phi/2} \end{aligned}$$

et les probabilités correspondantes

$$\begin{aligned} p(+\to[+,\hat{n}]) &= |\langle +,\hat{n}|+\rangle|^2 = \cos^2\theta/2 \\ p(+\to[-,\hat{n}]) &= |\langle -,\hat{n}|+\rangle|^2 = \sin^2\theta/2 \end{aligned}$$

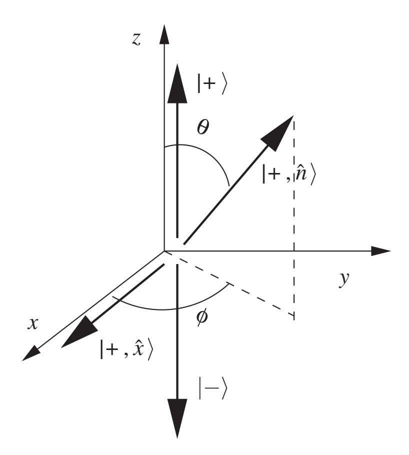


FIG. 3.13 – Représentation géométrique d'un état de spin $|\chi\rangle$. Le vecteur de Bloch $\vec{b} = \langle\vec{\sigma}\rangle$ est représenté par une flèche épaisse.

Nous avons obtenu l'essentiel des propriétés du spin 1/2, et ceci à partir des trois seules hypothèses, dont les deux premières découlent de l'invariance par rotation.

- La valeur moyenne $\langle\vec{S}\rangle$ se transforme comme un vecteur dans une rotation.
- Les valeurs propres de $\vec{S}\cdot\hat{n}$ sont indépendantes de $\hat{n}$.
- L'espace des états est de dimension deux.

Certaines de ces propriétés comme les relations de commutation (3.53) ou l'existence de matrices de rotation vont se transposer à un moment angulaire $\vec{J}$ quelconque (chapitre 9). Toutefois, d'autres propriétés sont spécifiques au spin 1/2 : par exemple c'est seulement dans ce cas que tout état de $\mathcal{H}$ peut s'écrire comme un vecteur propre de $\vec{J}\cdot\hat{n} = \vec{S}\cdot\hat{n}$.

3.2.5 Dynamique et évolution temporelle

Reprenons le problème du spin plongé dans un champ magnétique uniforme et constant $\vec{B}$, que nous supposerons orienté suivant l'axe des z. Notre étude classique du § 3.2.1 avait mis en évidence le phénomène de la précession de Larmor. En physique classique, l'énergie est un nombre

$$U = -\vec{\mu}\cdot\vec{B} = -\gamma\,\vec{s}\cdot\vec{B} = -\gamma\,s_z B = \omega s_z \tag{3.61}$$

où $\omega = -\gamma B$ est la fréquence de Larmor. En physique quantique, l'énergie devient un opérateur hermitien, que l'on appelle le *hamiltonien*, noté H, agissant dans l'espace des états. Comme cet espace est de dimension deux, le hamilto-

nien sera représenté par une matrice 2×2. Nous *admettrons*[23] qu'en mécanique quantique le hamiltonien conserve formellement l'expression (3.61), à condition de remplacer la quantité classique s_z par l'opérateur S_z, la projection suivant Oz de l'opérateur de spin $\vec{S}$

$$H = \omega S_z = \frac{\omega}{2}\,\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.62}$$

La deuxième forme de H donne sa représentation matricielle dans une base où S_z est diagonal. Les valeurs propres de H sont $+\hbar\omega/2$ et $-\hbar\omega/2$. Ce sont les deux valeurs possibles de l'énergie et les vecteurs propres correspondants sont bien sûr ceux de S_z : $|+\rangle$ et $|-\rangle$. Le schéma des niveaux d'énergie est donné sur la figure 3.14 pour $\omega > 0$, et les deux niveaux sont appelés *niveaux Zeeman* d'un spin 1/2 dans un champ $\vec{B}$.

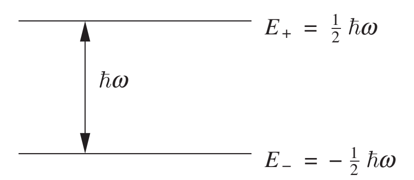


FIG. 3.14 – Spectre du hamiltonien (3.62).

Supposons qu'au temps $t = 0$, le spin se trouve dans l'état propre $|+, \hat{n}\rangle$. On peut alors se poser la question suivante : quel sera l'état de spin à un temps t ultérieur ? Pour répondre à cette question, nous avons besoin d'un postulat supplémentaire. Ce postulat, qui sera explicité avec plus de détails au chapitre suivant, stipule que le vecteur d'état $|\chi(t)\rangle$ au temps t se déduit du vecteur d'état au temps $t = 0$, $|\chi(t = 0)\rangle$, par

$$|\chi(t)\rangle = \exp\left(-\frac{\mathrm{i}Ht}{\hbar}\right)|\chi(0)\rangle \tag{3.63}$$

Cette loi d'évolution est particulièrement simple pour les vecteurs propres de H, appelés *états stationnaires*

$$|+\rangle \to \exp\left(-\frac{\mathrm{i}\omega t}{2}\right)|+\rangle \qquad |-\rangle \to \exp\left(\frac{\mathrm{i}\omega t}{2}\right)|-\rangle$$

Si $|\psi\rangle$ est un état arbitraire, la probabilité de trouver un état stationnaire dans $|\psi\rangle$ est *indépendante du temps* : par exemple

$$\left|\langle\psi\right| \exp\left(-\frac{\mathrm{i}Ht}{\hbar}\right)\left|+\rangle\right|^2 = |\langle\psi|+\rangle|^2$$

23. En dernier ressort, l'expression du hamiltonien trouve sa justification dans son accord avec l'expérience.

Supposons le spin orienté au temps $t = 0$ dans la direction $\hat{n}$

$$|\chi(0)\rangle = \cos\theta/2\exp(-\mathrm{i}\phi/2)|+\rangle + \sin\theta/2\exp(\mathrm{i}\phi/2)|-\rangle$$

ceci donnera au temps t

$$|\chi(t)\rangle = \cos\theta/2\exp[-\mathrm{i}(\phi+\omega t)/2]|+\rangle + \sin\theta/2\exp[\mathrm{i}(\phi+\omega t)/2]|-\rangle \quad (3.64)$$

Si au temps $t = 0$, le spin est orienté suivant la direction $\hat{n}$ définie par les angles θ et ϕ : $\langle\vec{S}\rangle = \frac{1}{2}\hbar\hat{n}$, au temps t le spin sera orienté dans la direction $(\theta, \phi+\omega t)$: le sens de rotation est le sens trigonométrique pour $q < 0$ et coïncide bien sûr avec celui du spin classique. La valeur moyenne du spin précesse autour de $\vec{B}$ avec la fréquence de Larmor.

La loi d'évolution (3.64) va nous permettre d'introduire une relation entre la dispersion ΔE sur l'énergie et le temps caractéristique d'évolution d'un système quantique, qui sera donnée sous la forme générale de *l'inégalité de Heisenberg temporelle* au § 4.2.4. Récrivons (3.64) en utilisant les notations c_+ et c_- pour les composantes de $|\chi(0)\rangle$ dans la base $\{|+\rangle, |-\rangle\}$

$$c_+ = \cos\theta/2\exp(-\mathrm{i}\phi/2) \qquad c_- = \sin\theta/2\exp(\mathrm{i}\phi/2)$$

et définissons les fréquences $\omega_\pm$

$$\omega_+ = \frac{E_+}{\hbar} = +\frac{1}{2}\,\omega \qquad \omega_- = \frac{E_-}{\hbar} = -\frac{1}{2}\,\omega$$

ce qui donne pour $|\chi(t)\rangle$

$$|\chi(t)\rangle = c_+\exp(-\mathrm{i}\omega_+ t)|+\rangle + c_-\exp(-\mathrm{i}\omega_- t)|-\rangle$$

Calculons la probabilité de trouver le vecteur d'état $|\chi(t)\rangle$ dans un état $|\psi\rangle$ arbitraire

$$\begin{aligned}|\langle\psi|\chi(t)\rangle|^2 &= |c_+|^2|\langle\psi|+\rangle|^2 + |c_-|^2|\langle\psi|-\rangle|^2 \\ &+ 2\mathrm{Re}\left[c_+^* c_-\exp[\mathrm{i}(\omega_+-\omega_-)t]\langle+|\psi\rangle\langle\psi|-\rangle\right]\end{aligned} \quad (3.65)$$

Les deux premiers termes de (3.65) sont indépendants du temps et le troisième oscille avec une fréquence

$$\omega_+ - \omega_- = \frac{E_+ - E_-}{\hbar} = \frac{\Delta E}{\hbar}$$

ΔE est la dispersion sur l'énergie : l'énergie du système n'a pas une valeur bien définie car le système passe d'un niveau à l'autre avec un temps caractéristique $\Delta t \simeq \hbar/\Delta E$, ce que l'on traduit par une *relation entre dispersion sur l'énergie et temps caractéristique d'évolution*

$$\Delta E\;\Delta t \simeq \hbar \quad (3.66)$$

Cette relation, que nous démontrerons sous la forme d'une inégalité par une méthode plus générale au § 4.2.4, est un exemple d'inégalité de Heisenberg temporelle.

3.3 Exercices

3.3.1 Polarisation elliptique et détermination de la polarisation

1. La polarisation d'une onde lumineuse est décrite par deux paramètres complexes

$$\lambda = \cos\theta\, \mathrm{e}^{\mathrm{i}\delta_x} \qquad \mu = \sin\theta\, \mathrm{e}^{\mathrm{i}\delta_y}$$

vérifiant $|\lambda|^2 + |\mu|^2 = 1$. De façon plus explicite, le champ électrique est

$$\begin{aligned} E_x(t) &= E_0 \cos\theta \cos(\omega t - \delta_x) = E_0 \mathrm{Re}\left(\cos\theta\, \mathrm{e}^{\mathrm{i}\delta_x} \mathrm{e}^{-\mathrm{i}\omega t}\right) \\ E_y(t) &= E_0 \sin\theta \cos(\omega t - \delta_y) = E_0 \mathrm{Re}\left(\sin\theta\, \mathrm{e}^{\mathrm{i}\delta_y} \mathrm{e}^{-\mathrm{i}\omega t}\right) \end{aligned}$$

Déterminer les axes de l'ellipse parcourue par l'extrémité du champ électrique et le sens de parcours.

2. On fait passer cette onde lumineuse à travers un polaroïd dont l'axe est parallèle à Ox. Montrer que la mesure de l'intensité à la sortie du polaroïd permet de déterminer θ.

3. On oriente maintenant le polaroïd suivant une direction faisant un angle de $\pi/4$ avec Ox. Quelle est la réduction d'intensité à la sortie du polaroïd ? Montrer que cette seconde mesure permet de déterminer la différence de phase $\delta = \delta_y - \delta_x$.

4. Vérifier que l'état $|\Phi_\perp\rangle$ (3.19) orthogonal à $|\Phi\rangle$

$$|\Phi_\perp\rangle = -\mu^*|x\rangle + \lambda^*|y\rangle$$

est arrêté par le polariseur linéaire du polariseur (λ, μ).

5. Vérifier que les propriétés physiques du polariseur (λ, μ) sont inchangées si l'on utilise la paramétrisation générale avec λ et μ complexes

$$\lambda = \cos\theta\, \mathrm{e}^{\mathrm{i}\eta_x} \qquad \mu = \sin\theta\, \mathrm{e}^{\mathrm{i}\eta_y}$$

avec $\eta = \eta_y - \eta_x$. Retrouver l'expression de $\mathcal{P}_\Phi$.

3.3.2 Une stratégie optimale pour Ève

Supposons que Ève analyse la polarisation du photon envoyé par Alice à l'aide d'un analyseur orienté $\updownarrow$. Si Alice oriente son polariseur $\updownarrow$, la probabilité pour Ève de trouver un photon $\updownarrow$, que l'on associera au résultat +1 d'une mesure, est de 100 %, mais elle est seulement de 50 % quand Alice utilise un polariseur $\nwarrow\!\!\searrow$. Sa probabilité de mesurer +1 quand Alice envoie aléatoirement $\updownarrow$ ou $\nwarrow\!\!\searrow$ est donc

$$\mathsf{p} = \frac{1}{2} + \frac{1}{2}\left(\frac{1}{2}\right) = \frac{3}{4}$$

Supposons que Ève oriente son analyseur suivant une direction faisant un angle ϕ avec Ox. Montrer que la probabilité $\mathsf{p}(\phi)$ pour Ève de mesurer $+1$ quand Alice envoie $+1$ est maintenant

$$\mathsf{p}(\phi) = \frac{1}{4}\,(2 + \cos 2\phi + \sin 2\phi)$$

Montrer que pour un choix optimal $\phi = \phi_0 = \pi/4$

$$\mathsf{p}(\phi_0) \simeq 0.854$$

une valeur plus élevée que précédemment. Pouvait-on prévoir sans calcul que la valeur optimale était $\phi = \phi_0 = \pi/8$?

2. Supposons qu'au lieu d'utiliser une base $|\pm\pi/4\rangle$, Alice et Bob utilisent une base $\{|\theta\rangle, |\theta_\perp\rangle\}$. Montrer que la probabilité d'erreur d'Ève est maintenant

$$\mathsf{p} = \frac{1}{2}\sin^2(2\theta) \qquad 0 < \theta < \frac{\pi}{4}$$

C'est l'utilisation d'une base complémentaire de la base $\{|x\rangle, |y\rangle\}$ qui maximise le taux d'erreur d'Ève.

3.3.3 Polarisation circulaire et opérateur de rotation pour les photons

1. Justifier les expressions suivantes pour les états $|D\rangle$ et $|G\rangle$ représentant des photons polarisés, respectivement à droite et à gauche

$$|D\rangle = -\frac{1}{\sqrt{2}}(|x\rangle + \mathrm{i}|y\rangle) \qquad |G\rangle = \frac{1}{\sqrt{2}}(|x\rangle - \mathrm{i}|y\rangle)$$

où $|x\rangle$ et $|y\rangle$ sont les vecteurs d'état de photons polarisés linéairement suivant Ox et Oy. Suggestion : quel est le champ électrique d'une onde lumineuse polarisée circulairement ? Écrire la forme matricielle des projecteurs $\mathcal{P}_D$ et $\mathcal{P}_G$ sur les états $|D\rangle$ et $|G\rangle$ dans la base $\{|x\rangle, |y\rangle\}$.

2. On définit les états $|\theta\rangle$ et $|\theta_\perp\rangle$ (3.11) représentant des photons polarisés linéairement suivant les directions faisant un angle θ avec respectivement Ox et Oy ainsi que

$$|D'\rangle = -\frac{1}{\sqrt{2}}(|\theta\rangle + \mathrm{i}|\theta_\perp\rangle) \qquad |G'\rangle = \frac{1}{\sqrt{2}}(|\theta\rangle - \mathrm{i}|\theta_\perp\rangle)$$

Comment $|D'\rangle$ et $G'\rangle$ sont-ils reliés à $|D\rangle$ et $|G\rangle$? Ces vecteurs d'état représentent-ils des états physiques différents de $|D\rangle$ et $|G\rangle$, et sinon pourquoi ?

3. On construit l'opérateur hermitien

$$\Sigma = \mathcal{P}_D - \mathcal{P}_G$$

Quelle est l'action de Σ sur les vecteurs $|D\rangle$ et $|G\rangle$? En déduire l'action de $\exp(-i\theta\Sigma)$ sur ces vecteurs.

4. Écrire la matrice représentative de Σ dans la base $\{|x\rangle, |y\rangle\}$. Montrer que $\Sigma^2 = I$ et retrouver $\exp(-i\theta\Sigma)$. En comparant avec la question **2**, donner l'interprétation physique de l'opérateur $\exp(-i\theta\Sigma)$.

3.3.4 Théorème de non-clonage quantique

On considère deux bases de polarisation linéaire pour des photons, la base $\{|V\rangle, |H\rangle\}$, avec $|V\rangle = |\updownarrow\rangle$ et $|H\rangle = |\leftrightarrow\rangle$, et la base $\{|A\rangle, |B\rangle\}$, avec

$$|A\rangle \equiv |\nwarrow\!\!\!\searrow\rangle = \frac{1}{\sqrt{2}}(|V\rangle + |H\rangle) \qquad |B\rangle \equiv |\swarrow\!\!\!\nearrow\rangle = \frac{1}{\sqrt{2}}(|V\rangle - |H\rangle)$$

On se propose de montrer que s'il existe un dispositif capable de recopier (cloner) les photons dans la base $\{|V\rangle, |H\rangle\}$, alors ce dispositif est incapable de le faire dans la base $\{|A\rangle, |B\rangle\}$. Supposons qu'il existe une photocopieuse quantique (une Q-photocopieuse) capable de recopier les états $|V\rangle$ et $|H\rangle$

$$|V \otimes X\rangle \Longrightarrow |V \otimes V\rangle \qquad |H \otimes X\rangle \Longrightarrow |H \otimes H\rangle$$

où X est la “page blanche” de la Q-photocopieuse et $\otimes$ désigne le produit tensoriel (§ 2.4). Quelle est l'action de la Q-photocopieuse sur l'état $|A \otimes X\rangle$? Montrer que l'on n'obtient pas l'état souhaité $|A \otimes A\rangle$.

3.3.5 Expérience à choix retardé

L'expérience à choix retardé de Jacques *et al.* [2007] utilise le montage schématisé sur la figure 3.15. La modification par rapport au montage de la figure 3.3 est la suivante : une lame demi-onde $\lambda/2$ fait tourner la polarisation de $\pi/4$, de telle sorte que les deux ondes ont la même phase à la sortie de l'interféromètre. En effet, les deux faisceaux ont un parcours identique dans l'indice ordinaire et dans l'indice extraordinaire. Un modulateur électro-optique MEO permet de faire tourner la polarisation de $\pi/4$. Le prisme polarisant PP sépare les photons polarisés verticalement des photons polarisés horizontalement, et un déphasage δ est introduit pour le faisceau supérieur.

1. On raisonne d'abord sur des ondes lumineuses. Lorsque le modulateur électro-optique n'est pas activé, montrer que le champ électrique entrant dans le prisme polarisant PP est de la forme

$$\vec{E} = \frac{1}{\sqrt{2}}\left(E^{(1)}\hat{y} + e^{i\delta}E^{(2)}\hat{x}\right)$$

où $E^{(1)}$ et $E^{(2)}$ ($|E^{(1)}| = |E^{(2)}| = E$) sont les amplitudes du champ provenant respectivement des trajets (1) et (2). Montrer que l'intensité lumineuse dans le détecteur D_1 est $|E^{(1)}|^2$ et celle dans le détecteur D_2 est $|E^{(2)}|^2$.

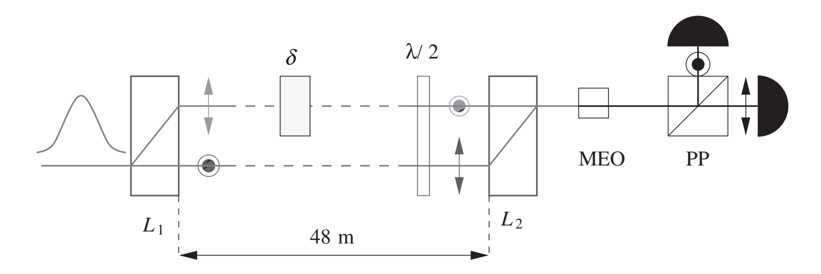


Fig. 3.15 – Schéma de l'expérience à choix retardé. Un photon unique polarisé entre dans l'interféromètre et rencontre une première lame biréfringente L_1 qui joue le rôle de lame séparatrice. Il "emprunte" un des deux trajets selon son état de polarisation verticale (trajet supérieur) ou horizontale (trajet inférieur). Une lame demi-onde λ échange les polarisations, de sorte que les deux trajets peuvent être recombinés par une seconde lame L_2. Le prisme polarisant PP permet de distinguer les états de polarisation et d'étiqueter le trajet suivi : il n'y a pas d'interférence, la polarisation joue comme un marqueur du trajet. Un modulateur électro-optique MEO peut faire tourner la polarisation de 45° et effacer l'information sur la polarisation. Si ce modulateur électro-optique est activé, on retrouve l'interférence. Un déphasage δ est introduit sur le trajet supérieur.

2. Le modulateur électro-optique fait tourner la polarisation de $\pi/4$

$$\hat{y} \to \frac{1}{\sqrt{2}}(\hat{y} + \hat{x}) \qquad \hat{x} \to \frac{1}{\sqrt{2}}(\hat{x} - \hat{y})$$

Montrer que les intensités lumineuses dans les détecteurs D_1 et D_2 sont proportionnelles à $(1 \pm \cos\delta)$. Adapter le raisonnement au cas des photons uniques.

3.3.6 Autres solutions de (3.45)

Dans l'espace des états de spin 1/2, la matrice unitaire $D^{(1/2)}(\theta, \psi)$ transforme l'état $|+\rangle$ en l'état $|+, \hat{n}\rangle$ où le vecteur unitaire $\hat{n}$ est donné par $\hat{n} = (\sin\theta\cos\psi, \sin\theta\sin\psi, \cos\theta)$. Si la rotation s'effectue autour de l'axe des z, alors $\theta = 0$ dans (3.58) et

$$D^{(1/2)}(\theta = 0, \psi) = U = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\psi/2} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\psi/2} \end{pmatrix}$$

Discuter l'action de U sur les états $|+\rangle$ et $|-\rangle$.

2. L'opérateur U peut être considéré comme un changement de base, où un opérateur A se transforme suivant (2.18) en

$$A \to A' = U^\dagger A U$$

Quels sont les opérateurs transformés de σ_x, σ_y et σ_z ?

3. Les conditions (3.45) ont pour solution soit (1) $\alpha - \alpha_x = \phi$ soit (2) $\alpha - \alpha_x = -\phi$. Montrer que dans le cas (1) σ_x et σ_y sont donnés par

$$\sigma_x = \begin{pmatrix} 0 & \mathrm{e}^{-\mathrm{i}\alpha_x} \\ \mathrm{e}^{-\mathrm{i}\alpha_x} & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i}\,\mathrm{e}^{-\mathrm{i}\alpha_x} \\ \mathrm{i}\,\mathrm{e}^{-\mathrm{i}\alpha_x} & 0 \end{pmatrix}$$

et que par rapport à la solution standard (3.47) cette solution corrrespond à une simple rotation des axes autour de Oz.

4. Montrer que si l'on choisit $\alpha - \alpha_x = -\phi$ la solution standard est

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & \mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix}$$

Quelle est l'interprétation de ce résultat ?

3.3.7 Décomposition d'une matrice 2×2

1. On introduit la notation :

$$\hat{\sigma}_0 = I \qquad \hat{\sigma}_i = \sigma_i, \quad i = 1, 2, 3$$

Montrer que si une matrice 2×2 A vérifie $\mathrm{Tr}\,(\hat{\sigma}_i A) = 0\ \forall i = 0, \ldots, 3$, alors $A = 0$.

2. Soit la matrice 2×2

$$A = \lambda_0 I + \sum_{i=1}^{3} \lambda_i \sigma_i = \sum_{i=0}^{3} \lambda_i \hat{\sigma}_i$$

Montrer que

$$\lambda_i = \frac{1}{2}\,\mathrm{Tr}\,(A\hat{\sigma}_i)$$

En déduire qu'une matrice 2×2 quelconque peut toujours s'écrire

$$A = \sum_{i=0}^{3} \lambda_i \hat{\sigma}_i$$

À quelle condition doivent obéir les coefficients λ_i lorsque A est hermitien, $A = A^\dagger$?

3.3.8 Exponentielles de matrices de Pauli

1. Montrer que toute matrice 2×2 unitaire et de déterminant unité U peut se mettre sous la forme (3.59). Suggestion : montrer que U est de la forme

$$\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}$$

et écrire $a = a_1 + \mathrm{i}a_2$, $b = b_1 + \mathrm{i}b_2$. Montrer que $a_1 = \cos\theta/2$.

2. Trouver deux matrices 2×2 A et B telles que

$$\mathrm{e}^A\,\mathrm{e}^B = \mathrm{e}^{(A+B)} \quad \text{avec} \quad [A, B] \neq 0$$

3.3.9 Tenseur ε_{ijk}

1. Montrer l'identité

$$\sum_k \varepsilon_{ijk}\varepsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$$

En déduire

$$\vec{a} \times (\vec{b} \times \vec{c}) = (a \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}$$

Que vaut

$$\sum_{jk} \varepsilon_{ijk}\varepsilon_{ljk} ?$$

2. On peut écrire la composante i du rotationnel d'un vecteur $\vec{A}$ comme

$$(\vec{\nabla} \times \vec{A})_i = \sum_{j,k} \varepsilon_{ijk}\partial_j A_k$$

avec $\partial_j = \partial/\partial x_j$. Montrer à partir de l'identité de la question **1** que

$$\vec{\nabla} \times \vec{\nabla} \times \vec{A} = \vec{\nabla}(\vec{\nabla} \cdot \vec{A}) - \nabla^2 \vec{A}$$

3.3.10 Mesures successives d'un spin 1/2

Un spin 1/2 est préparé dans un état up $|\hat{a}, +\rangle$ le long d'une direction $\hat{a}$. Le spin est ensuite mesuré le long d'une direction intermédiaire $\hat{b}$ et finalement suivant une direction $\hat{c}$. Montrer que si la dernière mesure donne le résultat $|\hat{c}, +\rangle$, alors la probabilité pour que la mesure intermédiaire ait donné le résultat $|\hat{b}, +\rangle$ est

$$\mathsf{p}(\hat{b}, +) = \frac{1}{1 + \tan^2 \theta_{ab}/2 \, \tan^2 \theta_{bc}/2}$$

où θ_{ab} et θ_{bc} sont les angles entre $(\hat{a}, \hat{b})$ et $(\hat{b}, \hat{c})$, respectivement. Suggestion : utiliser la loi de Bayes pour les probabilités conditionnelles sous la forme

$$\mathsf{p}(A|B)\mathsf{p}(B) = \mathsf{p}(B|A)\mathsf{p}(A)$$

avec $\mathsf{p}(A|B)$ = probabilité de A conditionnellement à B.

3.3.11 Rotation de 2π d'un spin 1/2

On reprend l'interféromètre à neutrons de l'exercice 1.6.6, le plan $ABDC$ étant horizontal. Un déphasage variable δ est obtenu en faisant passer les neutrons du faisceau I dans un champ magnétique uniforme et constant B sur une longueur l, le champ magnétique étant perpendiculaire au plan de la figure 3.16. Les neutrons sont supposés polarisés parallèlement au plan de la

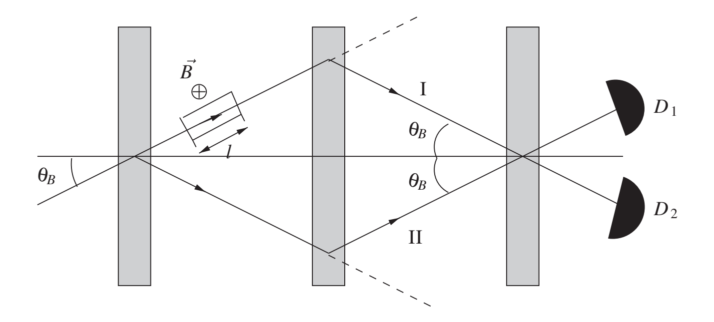


FIG. 3.16 – Mise en évidence expérimentale de la rotation de 2π d'un spin 1/2.

figure 3.16. Déterminer l'angle de la rotation subie par le spin du neutron à la sortie du champ magnétique en fonction de l, de la vitesse v (connue) du neutron et de son facteur gyromagnétique γ_n. Montrer que les taux de comptage par les détecteurs D_1 et D_2 dépendent sinusoïdalement de B. Montrer que l'on peut déduire de ces oscillations que le vecteur d'état de spin est multiplié par -1 dans une rotation de 2π (Werner *et al.* [1975]).

3.3.12 Diffusion de neutrons par un cristal : noyaux de spin 1/2

On reprend l'expérience décrite dans l'exercice 1.6.6 de diffraction de neutrons par un cristal en supposant que les noyaux atomiques ont un spin 1/2 (exemples : H^1, C^{13} F^{19}, etc.). On se limitera dans un premier temps (questions **1** et **2**) au cas où les neutrons ont un spin up (↑) et les noyaux un spin down (↓) : les neutrons et les noyaux sont polarisés. Dans ces conditions, il y a deux amplitudes de diffusion possibles car on peut montrer (§ 13.2.4) que la composante z du spin total est conservée dans la diffusion neutron-noyau. Ces deux amplitudes sont

- Une amplitude f_a où la diffusion se fait sans changement de l'état de spin

$$\text{neutron} \uparrow + \text{noyau} \downarrow \to \ \text{neutron} \uparrow + \text{noyau} \downarrow$$

- Une amplitude f_b où la diffusion s'effectue avec renversement du spin (spin flip)

$$\text{neutron} \uparrow + \text{noyau} \downarrow \to \ \text{neutron} \downarrow + \text{noyau} \uparrow$$

1. Montrer que dans le premier cas on retrouve les résultats de la diffusion sans spin.

2. Montrer que dans le second cas la diffraction disparaît et que la probabilité de diffusion est indépendante de $\vec{q}$.

3. En général les noyaux atomiques ne sont pas polarisés, c'est-à-dire qu'ils ont une chance sur deux d'avoir spin up et une chance sur deux d'avoir spin

down. On doit prendre en considération une troisième amplitude f_c correspondant à la diffusion

$$\text{neutron} \uparrow + \text{noyau} \uparrow \to \ \text{neutron} \uparrow + \text{noyau} \uparrow$$

Suivant la méthode utilisée dans l'exercice 1.6.7, introduisons un nombre α_i qui prend la valeur 0 si le noyau i a un spin up et la valeur 1 si ce noyau a un spin down. L'ensemble des $\{\alpha_i\}$ caractérise une configuration des spins dans le cristal. Montrer que l'amplitude de diffusion d'un neutron par le cristal dans la configuration $\{\alpha_i\}$ est

$$\sum_i (\alpha_i f_a + (1-\alpha_i) f_c)\, \mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}_i} + \sum_i \alpha_i f_b \mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}_i}$$

Que vaudrait l'intensité si la configuration $\{\alpha_i\}$ était fixée ? On prendra garde à additionner les probabilités pour des états finaux différents. On doit enfin prendre la moyenne sur les différentes configurations du cristal, le spin de chaque noyau étant supposé indépendant des autres spins. Si $\langle \bullet \rangle$ désigne la moyenne sur les configurations, montrer que

$$\langle \alpha_i \alpha_j \rangle = \frac{1}{4} + \frac{1}{4}\delta_{ij}$$

En déduire que la probabilité de diffusion est proportionnelle à

$$\mathcal{I} = \frac{1}{4}\,(f_a + f_c)^2 \sum_{i,j} \mathrm{e}^{\mathrm{i}\vec{q}\cdot(\vec{r}_i - \vec{r}_j)} + \frac{\mathcal{N}}{4}\,[(f_a - f_c)^2 + 2f_b^2]$$

où $\mathcal{N}$ est le nombre de noyaux. En réalité les trois amplitudes f_a, f_b et f_c ne sont pas indépendantes : on montre dans l'exercice 13.5.5 que

$$-f_a = \frac{1}{2}\,(a_t + a_s) \qquad -f_b = \frac{1}{2}\,(a_t - a_s) \qquad -f_c = a_t$$

où a_t et a_s sont les longueurs de diffusion dans les états triplet et singulet.

4. Que se passe-t-il si, comme c'est le cas courant en pratique, les neutrons ne sont pas polarisés ?

3.4 Bibliographie

La polarisation de la lumière et sa propagation dans les milieux anisotropes sont expliquées en détail dans May et Cazabat [1996], chapitres 19 et 20 ou Hecht [1987], chapitre 8. Comme complément à la discussion de la polarisation des photons, on pourra consulter Lévy-Leblond et Balibar [1984], chapitre 4 ou Baym [1969], chapitre 1. Un article de revue récent sur la cryptographie quantique, avec de nombreuses références aux travaux antérieurs, est celui de Gisin *et al.* [2002] ; une version grand public de la cryptographie quantique se trouve dans Bennett *et al.* [1992]. L'expérience de Stern-Gerlach est discutée par Feynman *et al.* [1965], volume III, chapitre 5, par Cohen-Tannoudji *et al.* [1973], chapitre IV, ou par Peres [1993], chapitre 1.

Chapitre 4

Postulats de la physique quantique

Nous allons énoncer dans ce chapitre les postulats de base de la physique quantique, en généralisant les résultats établis au chapitre précédent dans deux cas particuliers : la polarisation du photon et le spin 1/2. Au lieu d'être de dimension deux, l'espace des états sera *a priori* de dimension quelconque N, voire de dimension infinie. Les postulats tels qu'ils sont énoncés dans ce chapitre *fixent le cadre conceptuel général de la mécanique quantique*, et ne donnent pas directement les outils nécessaires pour résoudre des problèmes spécifiques. La résolution d'un problème de physique concret *suppose toujours une phase de modélisation*, où l'on simplifie le système à étudier, où l'on définit un cadre d'approximations, etc., et cette phase de modélisation s'appuie inévitablement sur des considérations plus ou moins heuristiques qui ne peuvent pas se déduire du cadre général de la physique quantique[1]. Le § 3.2.5 donne un exemple d'une telle démarche heuristique, conduisant à la solution d'un problème concret, celui du mouvement d'un spin 1/2 dans un champ magnétique.

Il est possible d'utiliser d'autres ensembles de postulats : par exemple une autre approche de la mécanique quantique consiste à énoncer des postulats sur les intégrales de chemin (§ 12.2.3). Comme c'est souvent le cas, une même théorie physique peut revêtir plusieurs habillages mathématiques différents. Enfin, il faut souligner que les postulats de la physique quantique soulèvent des problèmes épistémologiques difficiles, qui sont encore largement débattus aujourd'hui ; ils seront effleurés au § 11.4.7.

---

1. Cette démarche n'est pas fondamentalement différente de celle utilisée en physique classique. Par exemple, les trois lois de Newton fixent le cadre conceptuel de la mécanique classique, mais la solution d'un problème concret requiert toujours une phase de modélisation : simplification du problème posé, approximations pour les forces, etc.

4.1 Vecteurs d'état et propriétés physiques

4.1.1 Principe de superposition

Nous avons appris au chapitre 3 à caractériser l'état de polarisation d'un photon ou celui d'un spin 1/2 par un vecteur appartenant à un espace de Hilbert complexe, l'espace des états. Le postulat **I** généralise les notions de vecteur d'état et d'espace des états à tout système quantique.

Postulat I : espace des états

Les propriétés d'un système quantique sont entièrement définies par la donnée de son *vecteur d'état* $|\varphi\rangle$, qui fixe la représentation mathématique de l'état physique du système[2]. Le vecteur d'état est un élément d'un espace de Hilbert complexe $\mathcal{H}$ appelé *espace des états*. Il sera commode de choisir $|\varphi\rangle$ unitaire, c'est-à-dire de norme un : $||\varphi||^2 = \langle\varphi|\varphi\rangle = 1$.

Le fait qu'un état physique soit représenté par un vecteur implique sous certaines conditions le *principe de superposition*, caractéristique de la *linéarité* de la théorie : si $|\varphi\rangle$ et $|\chi\rangle$ sont des vecteurs de $\mathcal{H}$ représentant des états physiques, alors le vecteur unitaire de $\mathcal{H}$

$$|\psi\rangle = \frac{\lambda|\varphi\rangle + \mu|\chi\rangle}{||\lambda|\varphi\rangle + \mu|\chi\rangle||}$$

où λ et μ sont des nombres complexes, représente aussi un état physique.

Au chapitre précédent, nous avons défini les amplitudes de probabilité comme produits scalaires de vecteurs appartenant à l'espace des états. Par exemple, si $|\varphi\rangle$ représente l'état de polarisation linéaire suivant Ox d'un photon : $|\varphi\rangle = |x\rangle$, et $|\theta\rangle$ un état de polarisation linéaire suivant $\hat{n}_\theta$ (3.3) : $|\chi\rangle = |\theta\rangle$, l'amplitude de probabilité $a(x \to \theta) = \langle\theta|x\rangle = \cos\theta$. Nous avons également montré que le module carré de cette amplitude possède une interprétation physique remarquable : si l'on teste la polarisation en faisant passer le photon $|x\rangle$ à travers un analyseur linéaire d'orientation $\hat{n}_\theta$, on obtient une probabilité de transmission

$$\mathsf{p}(x \to \theta) = |a(x \to \theta)|^2 = |\langle\theta|x\rangle|^2 = \cos^2\theta$$

qui est la *probabilité pour le photon dans l'état* $|x\rangle$ *de passer le test* $|\theta\rangle$. Nous allons généraliser les notions d'amplitude de probabilité et de test en énonçant le postulat **II**.

---

2. Le point de vue de l'auteur est que le vecteur d'état décrit un système quantique *individuel*. Ce point de vue est loin d'être universellement partagé et le lecteur trouvera aisément d'autres interprétations, par exemple "le vecteur d'état décrit l'information disponible sur un système quantique", ou "le vecteur d'état n'est pas la propriété d'un système physique individuel, mais un protocole pour préparer un ensemble de tels états" ou encore "la mécanique quantique est un ensemble de règles permettant de calculer la probabilité d'un résultat expérimental" (§ 11.4.7). Cette diversité de points de vue n'a pas de conséquences sur l'utilisation *pratique* de la mécanique quantique.

Postulat II : amplitudes de probabilité et probabilités

Si $|\varphi\rangle$ est le vecteur représentant l'état du système et si $|\chi\rangle$ représente un autre état physique, il existe une *amplitude de probabilité* $a(\varphi \to \chi)$ *de trouver* $|\varphi\rangle$ *dans l'état* $|\chi\rangle$, qui est donnée par un produit scalaire sur $\mathcal{H}$: $a(\varphi \to \chi) = \langle\chi|\varphi\rangle$. La probabilité $\mathsf{p}(\varphi \to \chi)$ pour l'état $|\varphi\rangle$ de passer le test $|\chi\rangle$ s'obtient en prenant le module carré $|\langle\chi|\varphi\rangle|^2$ de cette amplitude[3],

$$\boxed{\mathsf{p}(\varphi \to \chi) = |a(\varphi \to \chi)|^2 = |\langle\chi|\varphi\rangle|^2} \tag{4.1}$$

ce qui constitue la *règle de Born.* Ajoutons quelques remarques pour compléter l'énoncé des deux premiers postulats.

- Sauf mention explicite du contraire, nous supposons les vecteurs d'état de norme unité. Si ce n'est pas le cas, il faut prendre garde à diviser par les normes. Par exemple la règle de Born (4.1) devient

$$\mathsf{p}(\varphi \to \chi) = \frac{|\langle\chi|\varphi\rangle|^2}{||\chi||^2||\varphi||^2}$$

- Les vecteurs $|\varphi\rangle$ et $|\varphi'\rangle = \exp(\mathrm{i}\beta)|\varphi\rangle$ représentent le même état physique. En effet on sait seulement mesurer des probabilités, et

$$|\langle\chi|\varphi\rangle|^2 = |\langle\chi|\varphi'\rangle|^2 \quad \forall\ |\chi\rangle \in \mathcal{H}$$

  Il n'est donc pas possible de distinguer entre $|\varphi\rangle$ et $|\varphi'\rangle$, qui diffèrent par un *facteur de phase*. En toute rigueur, un état physique est représenté par un *rayon*, ou vecteur à un facteur de phase près, de l'espace de Hilbert. L'ensemble des rayons forme un *espace de Hilbert projectif* : dans le cas du spin 1/2, cet espace peut être identifié à la sphère de Poincaré-Bloch. En revanche la superposition $\lambda|\varphi\rangle + \mu|\chi\rangle$ représente un état physique *différent* de $\lambda|\varphi'\rangle + \mu|\chi\rangle$! Décider entre les phases qui sont physiquement pertinentes et celles qui ne le sont pas, peut être un exercice délicat, comme le montre la phase de Berry (section 12.6).
- Le module du produit scalaire $S = |\langle\chi|\varphi\rangle|$ de deux états quantiques $|\varphi\rangle$ et $|\chi\rangle$ est une mesure quantitative de leur "discernabilité" dans le sens suivant : si l'on prépare une suite d'états $|\varphi\rangle$ et $|\chi\rangle$ avec respectivement une probabilité r et $(1-r)$, et si l'on essaie de distinguer ces deux états par une mesure, alors le taux maximum de succès $\mathsf{p}_{\max}$ de l'opération est donné par la *borne de Helstrom* (exercice 4.4.8), qui est une fonction de S

$$\mathsf{p}_{\max} = \frac{1}{2}\left(1 + \sqrt{1 - 4r(1-r)S^2}\right) = \frac{1}{2}\left(1 + \sqrt{1 - 4r(1-r)|\langle\varphi|\chi\rangle|^2}\right) \tag{4.2}$$

3. Afin que l'ordre des facteurs corresponde à celui du produit scalaire, il est parfois commode de noter les amplitudes de probabilité $a(\chi \leftarrow \varphi)$ et les probabilités $\mathsf{p}(\chi \leftarrow \varphi)$. On peut aussi observer qu'à défaut d'être intuitive, l'équation (4.1) est au moins cohérente : la probabilité de trouver l'état en lui-même est un et d'après l'inégalité de Schwarz, $0 \le |\langle\chi|\varphi\rangle|^2 \le 1$.

Le taux maximum de succès est de 100 % si $|\varphi\rangle$ et $|\chi\rangle$ sont orthogonaux ($S = 0$) : deux états orthogonaux sont parfaitement discernables.

- Nous nous limitons aux systèmes physiques qui sont appelés *cas purs*, ceux où l'information sur l'état physique est maximale. Dans le cas d'une information incomplète, on doit avoir recours au formalisme de l'opérateur statistique, qui sera exposé dans la section 11.1.
- Nous avons bien pris soin de préciser "système quantique", et non "particule" (quantique), qui en est un cas particulier. En effet, nous verrons au chapitre 11 que pour un système de deux ou plusieurs particules, il est en général impossible d'attribuer un vecteur d'état individuel à chacune des particules, et c'est seulement à *l'ensemble des particules*, c'est-à-dire à l'ensemble du système quantique, que l'on peut attribuer un vecteur d'état. Ce point sera développé et illustré dans la section 11.1.
- Il existe des restrictions au principe de superposition, appelées "règles de supersélection"[4], que nous n'aurons pas à prendre en compte dans ce livre.

4.1.2 Propriétés physiques et mesure

Au chapitre 3, nous avons montré qu'à la propriété physique "composante du spin suivant un axe $\hat{n}$", on pouvait faire correspondre un opérateur hermitien $\vec{S}\cdot\hat{n}$ agissant dans l'espace des états. Le postulat **III** généralise ce résultat à toute propriété physique.

Postulat III : propriétés physiques et opérateurs

À toute propriété physique $\mathcal{A}$ (énergie, position, impulsion, moment angulaire...) est associé un opérateur hermitien A agissant dans l'espace des états $\mathcal{H}$: A fixe la représentation mathématique de $\mathcal{A}$.

Afin de simplifier dans un premier temps la discussion qui va suivre, examinons le cas d'une propriété physique $\mathcal{A}$ représentée par un opérateur hermitien A dont les valeurs propres a_n sont non dégénérées : $A|n\rangle = a_n|n\rangle$. On peut alors écrire la décomposition spectrale

$$A = \sum_n |n\rangle a_n \langle n|$$

4. Il est généralement admis que l'on ne peut pas superposer un état de spin 1/2, $|\chi\rangle_{1/2}$ et un état de spin 1, $|\varphi\rangle_1$: cette impossibilité est un exemple de règle de supersélection. Comme nous l'avons vu au chapitre 3 (et cette observation sera généralisée au chapitre 9), le vecteur d'état d'une particule de spin 1/2 est multiplié par -1 dans une rotation de 2π, tandis que celui d'une particule de spin 1 est multiplié par $+1$. Dans une rotation de 2π qui ramène le système à sa situation initiale, si le vecteur d'état est de la forme $|\psi\rangle = \lambda|\varphi\rangle_1 + \mu|\chi\rangle_{1/2}$, ce vecteur d'état est transformé dans une rotation de 2π en $|\psi'\rangle = \lambda|\varphi\rangle_1 - \mu|\chi\rangle_{1/2} \neq |\psi\rangle$. Le fait que $|\chi\rangle_{1/2}$ soit transformé en $-|\chi\rangle_{1/2}$ ne pose aucun problème, car les deux vecteurs ne diffèrent que par un facteur de phase. Un autre exemple est la règle de supersélection sur la masse, dans le cas de l'invariance galiléenne. Pour un point de vue critique sur les règles de supersélection, voir Weinberg [1995], chapitre 2.

Si le sytème quantique est dans un état $|\varphi\rangle \equiv |n\rangle$, la valeur de l'opérateur A dans cet état est a_n : la propriété physique $\mathcal{A}$ prend la valeur numérique *exacte* a_n. Si $|\varphi\rangle$ n'est pas état (ou vecteur) propre de A, on sait d'après **II** que la probabilité $\mathsf{p}_n \equiv \mathsf{p}(a_n)$ de trouver $|\varphi\rangle$ dans $|n\rangle$, et donc de mesurer la valeur a_n de $\mathcal{A}$, est $\mathsf{p}_n = |\langle n|\varphi\rangle|^2$. Pour déterminer si le système quantique est dans l'état $|n\rangle$, $n = 1, \ldots, N$, on peut imaginer une généralisation de l'expérience de Stern-Gerlach avec N voies de sortie au lieu des deux voies $|+\rangle$ et $|-\rangle$ et un détecteur associé à chaque voie. Effectuons une série de tests sur des systèmes quantiques qui se trouvent tous dans l'état $|\varphi\rangle$. On dit que ces sytèmes ont été *préparés dans l'état* $|\varphi\rangle$: nous avons déjà rencontré la notion de préparation d'un système quantique dans le cas de la polarisation des photons, et nous y reviendrons ultérieurement. Si le nombre de tests $\mathcal{N}$ est très grand, on peut en déduire expérimentalement la *valeur moyenne de la propriété physique* $\mathcal{A}$ *dans l'état* $|\varphi\rangle$, notée $\langle A\rangle_\varphi$

$$\langle A\rangle_\varphi = \lim_{\mathcal{N}\to\infty} \frac{1}{\mathcal{N}} \sum_{p=1}^{\mathcal{N}} \mathcal{A}_p \tag{4.3}$$

où $\mathcal{A}_p$ est le résultat de la mesure n° p. $\mathcal{A}_p$ varie d'un test à l'autre, mais est toujours égal à l'une des valeurs propres a_n. Cette valeur moyenne est donnée en fonction de A et $|\varphi\rangle$ par

$$\langle A\rangle_\varphi = \sum_n \mathsf{p}_n a_n = \sum_n \langle\varphi|n\rangle a_n \langle n|\varphi\rangle = \langle\varphi|A|\varphi\rangle$$

Nous avons déjà rencontré un cas particulier de cette relation dans (3.38). Il n'est pas difficile de généraliser au cas des valeurs propres dégénérées. Si le système est dans un état $|\varphi\rangle$ quelconque, nous pouvons décomposer $|\varphi\rangle$ sur la base des vecteurs propres de A en utilisant la relation de fermeture (2.30)

$$|\varphi\rangle = \sum_{n,r} |n,r\rangle\langle n,r|\varphi\rangle = \sum_{n,r} c_{nr}|n,r\rangle$$

Pour trouver la probabilité $\mathsf{p}(a_n)$ d'observer la valeur propre a_n, il faut maintenant sommer sur l'indice r à n fixé toutes les probabilités de trouver $|\varphi\rangle$ dans l'un quelconque des états $|n,r\rangle$

$$\begin{aligned} \mathsf{p}(a_n) &= \sum_r |c_{nr}|^2 = \sum_r \langle\varphi|n,r\rangle\langle n,r|\varphi\rangle \\ &= \langle\varphi|\mathcal{P}_n|\varphi\rangle \end{aligned} \tag{4.4}$$

où $\mathcal{P}_n$ est le projecteur sur le sous-espace de la valeur propre a_n (*cf.* (2.29))

$$\mathcal{P}_n = \sum_r |n,r\rangle\langle n,r| \tag{4.5}$$

Comme ci-dessus, la répétition d'un grand nombre de mesures sur des systèmes quantiques préparés dans des conditions identiques permet d'obtenir la valeur moyenne $\langle A\rangle_\varphi$ de $\mathcal{A}$ dans l'état $|\varphi\rangle$

$$\langle A\rangle_\varphi = \sum_n a_n \mathsf{p}(a_n) = \sum_{n,r} \langle\varphi|n,r\rangle\, a_n\, \langle n,r|\varphi\rangle$$

soit en utilisant (2.31)

$$\boxed{\langle A\rangle_\varphi = \langle\varphi|A|\varphi\rangle} \qquad (4.6)$$

ce qui généralise le résultat précédent. Les opérateurs représentant des propriétés physiques sont souvent appelés "observables" dans la littérature. Nous éviterons cette terminologie qui n'a pas de véritable intérêt[5].

L'opérateur hermitien le plus simple est le projecteur sur un vecteur de $\mathcal{H}$, et faire passer un test $|\chi\rangle$ à un système quantique est équivalent à mesurer le projecteur $\mathcal{P}_\chi = |\chi\rangle\langle\chi|$: $\mathcal{P}_\chi$ prend la valeur un si le sytème passe le test $|\chi\rangle$ et zéro s'il échoue. Compte tenu de la décomposition spectrale d'un opérateur hermitien comme somme de projecteurs, on voit que *les notions de test et de mesure d'une propriété physique sont étroitement liées.* On mettra plutôt l'accent sur l'aspect "mesure" si l'on est intéressé par la valeur propre[6] de A, et plutôt sur l'aspect "test" si l'on est intéressé par la probabilité de trouver le système dans un état propre de A. Illustrons-le sur l'expérience de Stern-Gerlach du § 3.2.2. Dans l'interpétation "mesure du spin", l'appareil de Stern-Gerlach mesure la composante z du spin en déviant les atomes d'argent vers le haut ou vers le bas, et la détection de l'atome sur un écran à la sortie de l'appareil permet de distinguer entre les valeurs $+\hbar/2$ et $-\hbar/2$ de la propriété physique $\mathcal{S}_z$, composante du spin suivant l'axe Oz. On peut, de façon équivalente, dire que l'on fait passer aux atomes les tests $|+\rangle$ et $|-\rangle$. La probabilité de déviation vers le haut (resp. bas) est $|\langle+|\varphi\rangle|^2$ (resp. $|\langle-|\varphi\rangle|^2$).

Cependant, les mesures, ou les tests, décrits au § 3.2.2 présentent un inconvénient : la mesure n'est achevée que lorsque les atomes sont absorbés sur l'écran et ils ne sont plus disponibles pour des expériences ultérieures. Dans une *mesure idéale (ou test idéal)*, on suppose que le système physique n'est pas détruit par la mesure[7], et que si avant la mesure de $\mathcal{A}$ le vecteur d'état

5. Cette terminologie remonte à l'article fondateur de Heisenberg, dont est extraite la citation suivante : "Cet article a pour objet d'établir que la théorie quantique est fondée exclusivement sur des relations entre quantités qui sont en principe observables". Se restreindre à une telle approche est une vision étroite de la physique, que Heisenberg lui-même n'a pas respectée dans sa pratique !

6. On peut donner une formulation "mesure" au test de polarisation d'un photon, par exemple dans la base $\{|x\rangle, |y\rangle\}$, en introduisant la propriété physique $\mathcal{A}_x$ représentée par l'opérateur

$$A_x = |x\rangle\langle x| - |y\rangle\langle y|$$

qui prend la valeur $+1$ si le photon est polarisé suivant Ox et -1 s'il est polarisé suivant Oy.

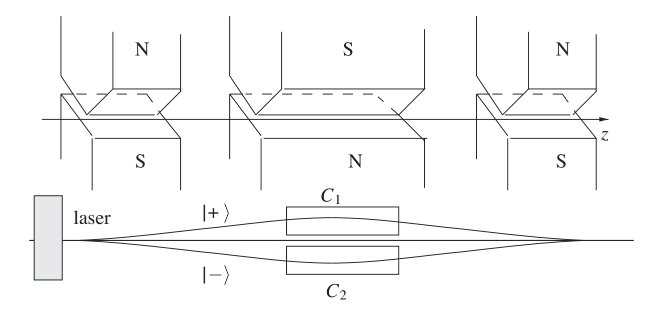


FIG. 4.1 – Mesure idéale du spin.

était $|\varphi\rangle = \sum_n c_n|n\rangle$, le système *après* une mesure donnant le résultat a_n doit être dans l'état $|\varphi_n\rangle$. On pourrait imaginer une mesure idéale[8] (tout à fait théorique !) du spin grâce à un filtre de Stern-Gerlach modifié s'inspirant du dispositif décrit au § 1.4.4. Prenant comme base de départ le filtre de la figure 3.9, on illumine l'atome entrant dans le filtre par un faisceau laser convenablement accordé qui induit une transition dans un niveau excité de l'atome. Lorsqu'elles ont leur séparation maximale à l'intérieur du filtre, les deux trajectoires passent dans deux cavités résonantes distinctes où l'atome revient dans son état fondamental en émettant un photon avec une probabilité voisine de 100 % (figure 4.1). Ce photon est détecté dans l'une des deux cavités, et il est ainsi possible d'étiqueter la trajectoire à l'intérieur du filtre, *sans perturber en quoi que ce soit l'état de spin.* Cette mesure entraîne une profonde modification dans la description de l'état de spin. Si, par exemple, l'état du spin à l'entrée du filtre est l'état propre $|+,\hat{x}\rangle$ de S_x, *en l'absence de mesure*, les deux trajectoires conservent la propriété de cohérence. Elles peuvent être recombinées à la sortie du filtre pour reconstruire l'état $|+,\hat{x}\rangle$: le filtre contient une superposition *cohérente* d'états propres de S_z, $|+\rangle$ et $|-\rangle$, avec une amplitude $1/\sqrt{2}$

$$|+,\hat{x}\rangle = \frac{1}{\sqrt{2}}\big(|+\rangle + |-\rangle\big)$$

Au contraire, lorsqu'une mesure est effectuée, le spin est projeté sur l'un des états $|+\rangle$ ou $|-\rangle$ avec une probabilité de 50 %, et il est impossible de revenir en arrière et de reconstruire l'état $|+,\hat{x}\rangle$. Si l'on n'observe pas le résultat de la mesure, on peut interpréter celle-ci comme transformant la superposition cohérente $|+,\hat{x}\rangle$ en un *ensemble statistique classique* de 50 % de spins up et 50 % de spins down.

7. Si l'on peut répéter plusieurs fois une même mesure idéale, on a alors une “mesure quantique sans démolition” ou mesure QND (Quantum Non Demolition). Voir par exemple Caves *et al.* [1980] ou Braginsky *et al.* [1980] et l'exercice 17.5.7 pour une réalisation concrète d'une telle mesure.

8. Une autre expérience théorique a été proposée par Scully *et al.* [1989].

Si la mesure de $\mathcal{S}_z$ a donné le résultat $+\hbar/2$ et si on répète cette mesure, on constate que le résultat est toujours $+\hbar/2$: immédiatement après une mesure de $\mathcal{S}_z$ qui a donné le résultat $+\hbar/2$, le spin est dans l'état $|+\rangle$. Cette observation pose la question de deux mesures idéales consécutives. L'analyse de cette situation requiert un formalisme qui sera établi au chapitre 11, et elle sera effectuée au § 11.4.6. Le résultat de cette analyse est le suivant : tout se passe comme si à l'issue d'une première mesure, lorsqu'un système initialement dans l'état $|\varphi\rangle$ a passé avec succès le test $|\chi\rangle$, le système était projeté après le test dans l'état $|\chi\rangle$

$$|\varphi\rangle \to \frac{\mathcal{P}_\chi|\varphi\rangle}{||\mathcal{P}_\chi|\varphi\rangle||}$$

où le dénominateur sert à normaliser l'état. Ce résultat est présenté dans la plupart des manuels comme un postulat supplémentaire de la mécanique quantique, le "postulat de réduction du paquet d'ondes" (RPO), énoncé sous la forme :

Postulat RPO. Si le système était initialement dans l'état $|\varphi\rangle$, et si le résultat de la mesure de $\mathcal{A}$ est a_n, alors immédiatement après la mesure, le système se trouve dans l'état projeté sur le sous-espace de la valeur propre a_n

$$|\varphi\rangle \to |\psi\rangle = \frac{\mathcal{P}_n|\varphi\rangle}{(\langle\varphi|\mathcal{P}_n|\varphi\rangle)^{1/2}} \tag{4.7}$$

Le vecteur $|\psi\rangle$ dans (4.7) est bien normalisé à l'unité car

$$||\mathcal{P}_n|\varphi\rangle||^2 = \langle\varphi|\mathcal{P}_n^\dagger\mathcal{P}_n|\varphi\rangle = \langle\varphi|\mathcal{P}_n|\varphi\rangle$$

compte tenu des propriétés des projecteurs. L'énoncé de ce pseudo-postulat appelle quelques remarques. Il ne faut surtout pas imaginer que la transformation (4.7) correspond à un processus physique réel. Elle n'a de sens que si l'on s'intéresse à une évolution effective, qui fait abstraction de l'appareil de mesure pour se focaliser uniquement sur le système. En fait, c'est une simple commodité d'écriture dans l'espace de Hilbert des états du système, qui isole artificiellement le système de l'appareil de mesure et de son environnement. De plus (*cf.* § 11.4.6), elle n'est valable que pour une mesure idéale[9].

Nous avons donné ci-dessus un exemple de mesure idéale pour le spin d'un atome d'argent (figure 4.1). À la sortie du filtre, on sait dans quel état de spin

9. Il aurait peut-être été préférable de passer purement et simplement sous silence ce "postulat RPO". Nous l'avons énoncé afin que le lecteur puisse faire le lien avec les exposés traditionnels. En fait, ce postulat est indissociable du statut ontologique du vecteur d'état. Par exemple, dans l'interprétation statistique de Ballentine [1998], chapitre 9, le formalisme de la mécanique quantique ne s'applique qu'à des ensembles de systèmes préparés de façon identique, et non à des systèmes individuels, contrairement à ce que nous avons admis (voir la note 2). Il est clair que ce postulat n'a simplement aucun sens dans l'interpétation statistique. Pour une analyse critique de la "réduction du paquet d'ondes", voir Ballentine [1990] ou Cohen-Tannoudji [1990].

se trouve l'atome qui est maintenant disponible pour des tests ultérieurs. Une répétition de la mesure de $\mathcal{S}_z$ redonnera $+\hbar/2$ pour les atomes qui ont émis un photon dans C_1 et $-\hbar/2$ pour ceux qui ont émis un photon dans C_2. Il faut remarquer que la mesure idéale se présente rarement en pratique. En général, la détection détruit le système observé : un exemple déjà mentionné de mesure destructrice[10] est la détection d'un photon par un photodétecteur D_x ou D_y sur figure 3.2. Un autre exemple de mesure non idéale est la détermination de l'impulsion d'une particule par collision élastique avec une seconde particule d'impulsion connue, en utilisant la conservation de l'énergie-impulsion. Après la collision, la première particule n'est pas détruite, mais elle ne se trouve plus dans l'état d'impulsion que l'on a mesurée. Le concept de mesure idéale est utile pour la discussion de la mesure en physique quantique, mais en pratique la mesure idéale est l'exception, et non la règle !

Lorsque l'on cherche à déterminer complètement le vecteur d'état $|\varphi\rangle$ d'un système physique, il peut arriver que la mesure idéale d'une propriété physique $\mathcal{A}$ donne le résultat a, la valeur propre a de A étant non dégénérée. Immédiatement après la mesure, le vecteur d'état est alors le vecteur propre $|a\rangle$ de A. Si la valeur propre est dégénérée, il faut trouver une seconde propriété physique $\mathcal{B}$ compatible avec $\mathcal{A}$: $[A, B] = 0$. Dans ce cas, il est possible que la donnée des valeurs propres a et b spécifie entièrement le vecteur d'état. Si ce n'est pas encore le cas, il faudra trouver une troisième propriété physique $\mathcal{C}$ compatible avec $\mathcal{A}$ et $\mathcal{B}$, etc. Lorsque la donnée des valeurs propres $\{a, b, c \dots\}$ des opérateurs compatibles $\{A, B, C \dots\}$ spécifie entièrement le vecteur d'état on dira, en suivant la terminologie introduite au § 2.3.3, que ces opérateurs (ou les propriétés physiques qu'ils représentent) forment un *ensemble complet d'opérateurs (ou de propriétés physiques) compatibles*. La mesure simultanée d'un système complet de propriétés physiques compatibles $\{\mathcal{A}, \mathcal{B}, \mathcal{C} \dots\}$ constitue un *test maximal* du vecteur d'état. Si l'espace des états est de dimension N, *un test maximal doit avoir N résultats différents possibles.* Lorsque l'on a réalisé un test maximal sur un système quantique, on connaît exactement son vecteur d'état, et on a donc *préparé* le système quantique dans un état déterminé : on a effectué l'étape de préparation du système. Cependant, la préparation n'implique pas nécessairement une mesure : par exemple le filtre de gauche de la figure 3.10 prépare le spin 1/2 dans l'état $|+\rangle$ en éliminant tous les spins qui se trouvent dans $|-\rangle$, sans qu'une mesure du spin ne soit effectuée.

Pour fixer les idées, supposons que la donnée de deux valeurs propres a_r et b_s de deux opérateurs compatibles A et B spécifie entièrement un vecteur $|r, s\rangle$ de $\mathcal{H}$

$$A|r, s\rangle = a_r|r, s\rangle \qquad B|r, s\rangle = b_s|r, s\rangle$$

La mesure simultanée des propriétés physiques $\mathcal{A}$ et $\mathcal{B}$ est alors un test maximal et les N résultats possibles sont étiquetés par le couple (r, s). Un exemple

10. On sait maintenant effectuer des mesures non destructrices sur un photon : voir Gleyzes *et al.* [2007] et l'exercice 17.5.7.

d'appareil effectuant un test maximal est l'appareil de Stern-Gerlach de la figure 3.8 : cet appareil sépare les états de spin $|+\rangle$ et $|-\rangle$, qui donnent deux taches différentes sur l'écran, l'espace des états étant de dimension 2 : $N = 2$. La mesure de $\mathcal{A}$ et $\mathcal{B}$ permet de préparer le système dans l'état $|r, s\rangle$, en sélectionnant les systèmes qui ont donné le résultat (a_r, b_s). Si les systèmes quantiques sélectionnés dans l'état $|r, s\rangle$ sont à nouveau soumis à une mesure simultanée de $\mathcal{A}$ et $\mathcal{B}$, le résultat de cette nouvelle mesure sera (a_r, b_s) avec une probabilité de 100 %. Lorsqu'un système physique est décrit par un vecteur d'état, il doit exister, au moins en principe, un test maximal dont un des résultats possibles a une probabilité de 100 % : pour un spin 1/2 dans l'état $|+\rangle$, un tel test maximal est celui effectué avec un appareil de Stern-Gerlach dont le champ magnétique est parallèle à Oz.

Il est aussi instructif d'examiner le cas d'une propriété physique $\mathcal{A}$ compatible avec $\mathcal{B}$ et $\mathcal{C}$: $[A, B] = [A, C] = 0$, alors que $\mathcal{B}$ et $\mathcal{C}$ sont incompatibles : $[B, C] \neq 0$. Dans ce cas, le résultat de la mesure de $\mathcal{A}$ dépend de ce qu'on la mesure simultanément à $\mathcal{B}$ ou à $\mathcal{C}$. Cette propriété est appelée *contextualité*, et un exemple en sera donné au § 11.3.4.

Le lecteur se sera rendu compte que la mesure en physique quantique est fondamentalement différente de la mesure en physique classique. En physique classique, la mesure révèle une propriété *préexistante* du système physique testé. Si une voiture roule à 180 km/h sur l'autoroute, la mesure de sa vitesse par un radar détermine une propriété préexistante à la mesure, ce qui donne au gendarme la légitimité pour verbaliser. Au contraire, la mesure de la composante $\mathcal{S}_x$ d'un spin 1/2 dans l'état $|+\rangle$ ne révèle pas une valeur de $\mathcal{S}_x$ préexistante. Si l'on reprend l'analogie de la voiture, on pourrait imaginer que celle-ci soit dans un état de superposition linéaire de deux vitesses[11], par exemple

$$|v\rangle = \sqrt{\frac{1}{3}}\,|120 \text{ km/h}\rangle + \sqrt{\frac{2}{3}}\,|180\,\text{km/h}\rangle$$

Le gendarme mesurera une vitesse de 120 km/h avec une probabilité de 1/3 et une vitesse de 180 km/h avec une probabilité de 2/3, mais il serait erroné de penser que la voiture roulait à l'une des deux vitesses avant la mesure. La dispersion des résultats de la mesure d'une propriété physique quantique est parfois attribuée à la “perturbation incontrôlable due à la mesure”, mais la valeur de cette propriété ne préexiste pas à la mesure, et ce qui n'existe pas ne peut pas être perturbé.

4.1.3 Inégalités de Heisenberg II

Nous avons introduit au chapitre précédent la notion de propriétés physiques incompatibles. Nous allons revenir de façon plus quantitative sur ce

11. Bien sûr, on ne sait pas réaliser un tel état avec une voiture, mais on sait très bien fabriquer une particule élémentaire ou un atome dans un état de superposition linéaire de deux vitesses.

concept et ses conséquences pour la mesure. Deux propriétés physiques $\mathcal{A}$ et $\mathcal{B}$ sont *incompatibles* si le commutateur des opérateurs A et B qui les représentent est non nul : $[A, B] \neq 0$. Supposons qu'une première mesure de A ait donné un résultat a et projeté le vecteur d'état initial sur un vecteur propre $|a\rangle$ de A : $A|a\rangle = a|a\rangle$. Si l'on effectue une mesure de $\mathcal{B}$ immédiatement après celle de $\mathcal{A}$, en général le vecteur $|a\rangle$ ne sera pas vecteur propre de B et le résultat de la mesure ne sera connu qu'avec une certaine probabilité. Par exemple, si b est une valeur propre simple de B correspondant au vecteur propre $|b\rangle$: $B|b\rangle = b|b\rangle$, la probabilité de mesurer b sera $\mathsf{p}(a \to b) = |\langle b|a\rangle|^2$. En général, il ne sera pas possible de trouver des états où les valeurs de $\mathcal{A}$ et $\mathcal{B}$ soient toutes deux exactement connues. Nous allons établir un résultat important sur les dispersions (ou écarts types) des mesures effectuées à partir d'un état initial $|\varphi\rangle$ arbitraire. Comme en théorie des probabilités ordinaire, nous définissons les *dispersions*, ou *écarts types*, $\Delta_\varphi A$ et $\Delta_\varphi B$ *dans l'état* $|\varphi\rangle$ par

$$\begin{aligned} (\Delta_\varphi A)^2 &= \langle A^2\rangle_\varphi - (\langle A\rangle_\varphi)^2 = \langle (A - \langle A\rangle_\varphi I)^2\rangle_\varphi \\ (\Delta_\varphi B)^2 &= \langle B^2\rangle_\varphi - (\langle B\rangle_\varphi)^2 = \langle (B - \langle B\rangle_\varphi I)^2\rangle_\varphi \end{aligned} \tag{4.8}$$

Le commutateur de A et de B est de la forme iC, où C est un opérateur hermitien ; en effet

$$[A, B]^\dagger = [B^\dagger, A^\dagger] = [B, A] = -[A, B]$$

Nous pouvons donc écrire

$$[A, B] = \mathrm{i}C \qquad\qquad C = C^\dagger \tag{4.9}$$

Définissons les opérateurs hermitiens de valeur moyenne nulle (*a priori* uniquement dans l'état $|\varphi\rangle$)

$$A_0 = A - \langle A\rangle_\varphi I \qquad\qquad B_0 = B - \langle B\rangle_\varphi I$$

et dont le commutateur est aussi iC : $[A_0, B_0] = \mathrm{i}C$, car $\langle A\rangle_\varphi$ et $\langle B\rangle_\varphi$ sont des nombres. La norme au carré du vecteur

$$(A_0 + \mathrm{i}\lambda B_0)|\varphi\rangle$$

où λ est choisi réel, doit être positive

$$\begin{aligned} ||(A_0 + \mathrm{i}\lambda B_0)|\varphi\rangle||^2 &= ||A_0|\varphi\rangle||^2 + \mathrm{i}\lambda\langle\varphi|A_0B_0|\varphi\rangle - \mathrm{i}\lambda\langle\varphi|B_0A_0|\varphi\rangle \\ &\quad + \lambda^2||B_0|\varphi\rangle||^2 \\ &= \langle A_0^2\rangle_\varphi - \lambda\langle C\rangle_\varphi + \lambda^2\langle B_0^2\rangle_\varphi \geq 0 \end{aligned}$$

Le polynôme de degré deux en λ doit être positif quel que soit λ, ce qui implique

$$\langle C\rangle_\varphi^2 - 4\langle A_0^2\rangle_\varphi\langle B_0^2\rangle_\varphi \leq 0$$

Ceci démontre *l'inégalité de Heisenberg*

$$\boxed{(\Delta_\varphi A)\,(\Delta_\varphi B) \geq \frac{1}{2}\left|\langle C\rangle_\varphi\right|} \tag{4.10}$$

C'est la relation souhaitée donnant les dispersions sur les mesures de $\mathcal{A}$ et $\mathcal{B}$: le produit des dispersions sur les mesures est supérieur ou égal à la moitié du module de la valeur moyenne du commutateur de A et B. Il est facile de montrer (exercice 4.4.1) qu'une condition nécessaire et suffisante pour que $\Delta_\varphi A = 0$ est que $|\varphi\rangle$ soit vecteur propre de A. Dans un espace vectoriel de *dimension finie*, on a alors $\langle C\rangle_\varphi = 0$. Insistons sur l'interprétation correcte de (4.10) : en effectuant comme en (4.3) un grand nombre de mesures de $\mathcal{A}$, un grand nombre de mesures de $\mathcal{B}$ et un grand nombre de mesures de $\mathcal{C}$ sur des systèmes tous préparés dans le même état $|\varphi\rangle$, on pourra en déduire avec une bonne précision les dispersions $\Delta_\varphi A$ et $\Delta_\varphi B$ ainsi que la valeur moyenne $\langle C\rangle_\varphi$, qui obéiront alors à (4.10). Nous soulignons que $\mathcal{A}$, $\mathcal{B}$ et $\mathcal{C}$ sont bien sûr mesurés sans des expériences *différentes* : ils ne peuvent pas être mesurés simultanément si A, B et C ne commutent pas. De plus, $\Delta_\varphi A$ et $\Delta_\varphi B$ ne sont en rien liés aux erreurs de mesure. Si, par exemple, δA est la résolution expérimentale pour la mesure de $\mathcal{A}$, nous devons avoir $\delta A \ll \Delta_\varphi A$ pour une détermination précise de la dispersion. L'erreur sur $\langle A\rangle$ est contrôlée par la résolution expérimentale, et pas du tout par $\Delta_\varphi A$: rien n'empêche $\langle A\rangle_\varphi$ d'être déterminé avec une précision bien meilleure que $\Delta_\varphi A$.

4.2 Évolution temporelle

4.2.1 Équation d'évolution

Jusqu'à présent, nous avons considéré le système physique à un instant donné, ou pendant l'intervalle de temps supposé infiniment court d'une mesure. Nous allons maintenant prendre en considération l'évolution temporelle du vecteur d'état, auquel nous donnerons une dépendance explicite $|\varphi(t)\rangle$ par rapport au temps t.

Postulat IV : équation d'évolution
L'évolution temporelle du vecteur d'état $|\varphi(t)\rangle$ d'un système quantique *fermé* est régie par l'équation d'évolution

$$\boxed{\mathrm{i}\hbar\frac{\mathrm{d}|\varphi(t)\rangle}{\mathrm{d}t} = H(t)|\varphi(t)\rangle} \tag{4.11}$$

L'opérateur hermitien $H(t)$ est appelé *hamiltonien*. En écrivant (4.11), nous avons admis la nécessité d'un postulat supplémentaire indépendant pour l'évolution temporelle. En réalité, on peut *déduire* ce postulat moyennant une hypothèse plus faible, celle d'une évolution réversible, ainsi que nous le verrons à la fin du § 4.2.2.

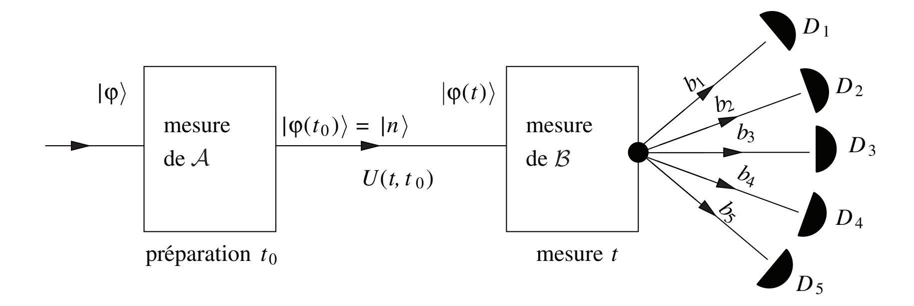


FIG. 4.2 – Préparation et mesure. La mesure de $\mathcal{A}$ au temps t_0 donne le résultat a_n et sélectionne l'état $|n\rangle$, $A|n\rangle = a_n|n\rangle$: c'est la préparation du système. L'évolution (4.14) entre t_0 et t se traduit par $|\varphi(t)\rangle = U(t, t_0)|\varphi(t_0\rangle$. Une mesure de $\mathcal{B}$ est ensuite effectuée au temps t : $B|m\rangle = b_m|m\rangle$, $m = 1, \ldots, M$, avec $M = 5$ sur la figure. Un filtre de Stern-Gerlach généralisé avec M voies de sortie réalise un test maximal. La probabilité de déclencher le détecteur D_m est $\mathsf{p}(b_m) = |\langle m|\varphi(t)\rangle|^2 = |\langle m|U(t, t_0)|n\rangle|^2$.

Il nous faut énoncer de façon précise les conditions de validité de (4.11). Cette équation est valable pour un *système quantique fermé*, ce qui doit se comprendre comme suit : *le système quantique considéré ne doit pas être une partie d'un système quantique plus grand*, une situation qui sera examinée en détail au chapitre 18. Cependant, (4.11) est valable si le système quantique interagit avec un système *classique*[12], ce qui veut dire qu'il n'est pas nécessairement isolé. L'équation (4.11), par exemple, est valable pour un spin 1/2 soumis à un champ magnétique variable (section 5.2), ou pour un atome soumis au champ électromagnétique d'une onde laser (sections 15.3.1 à 15.3.3), mais pas pour un atome interagissant avec un champ électromagnétique quantifié (§ 17.3.4). Dans ce dernier cas, l'évolution temporelle du vecteur d'état de l'atome, ou plus exactement celle de son opérateur statistique, n'est pas régie par un hamiltonien. Une évolution hamiltonienne n'est valable que pour l'ensemble *atome + champ*.

L'opérateur H a les dimensions d'une énergie, et nous identifierons effectivement H comme étant l'opérateur hermitien représentant la propriété physique énergie. L'équation (4.11) est du premier ordre par rapport au temps, et *l'évolution est déterministe* : étant donné une condition initiale $|\varphi(t_0)\rangle$ au temps $t = t_0$ pour le vecteur d'état, l'évolution (4.11) détermine $|\varphi(t)\rangle$ à tout temps ultérieur $t > t_0$, pourvu bien sûr que le hamiltonien soit connu. En fait, la restriction à $t > t_0$ n'est pas nécessaire : l'évolution (4.11) est réversible et on peut parfaitement "remonter le temps". Le schéma d'une expérience typique est donné sur la figure 4.2 : le système est préparé au temps $t = t_0$ par

12. La raison pour laquelle ceci est possible est subtile. L'explication est donnée dans l'exercice 17.5.11 : un champ classique ne s'intrique pas avec le système quantique (voir le chapitre 11 pour la définition de l'intrication).

la mesure d'un ensemble de propriétés physiques compatibles, qui détermine le vecteur d'état $|\varphi(t_0)\rangle$ (mais ainsi que nous l'avons déjà noté, la phase de préparation ne correspond pas nécessairement à un test maximal). Le vecteur d'état évolue ensuite jusqu'au temps t en suivant (4.11), et une seconde mesure d'une ou d'un ensemble de propriétés physiques (soit identiques à celles de la première mesure, soit différentes) est effectuée au temps t. Cette seconde mesure permet de déterminer totalement ou en partie $|\varphi(t)\rangle$, et par exemple de remonter aux propriétés de H. Pour que (4.11) soit valable entre les deux mesures, il est bien sûr nécessaire que le système quantique reste fermé dans l'intervalle de temps correspondant.

La (nécessaire) conservation de la norme du vecteur d'état est assurée par l'hermiticité de H. En effet,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\,||\varphi(t)||^2 &= \frac{\mathrm{d}}{\mathrm{d}t}\,\langle\varphi(t)|\varphi(t)\rangle \\ &= \langle\varphi(t)|\Big(\frac{1}{\mathrm{i}\hbar}\,H\Big)^\dagger|\varphi(t)\rangle + \langle\varphi(t)|\Big(\frac{1}{\mathrm{i}\hbar}\,H\Big)|\varphi(t)\rangle \\ &= \frac{1}{\mathrm{i}\hbar}\,\langle\varphi(t)|(H - H^\dagger)|\varphi(t)\rangle = 0 \qquad (4.12)\end{aligned}$$

car $H = H^\dagger$. Si l'on décompose $|\varphi(t)\rangle$ sur une base $|n, r\rangle$

$$|\varphi(t)\rangle = \sum_{n,r} |n,r\rangle\langle n,r|\varphi(t)\rangle = \sum_{n,r} c_{nr}(t)|n,r\rangle$$

les composantes $c_{nr}(t)$ obéissent à

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\sum_{n,r}|c_{nr}(t)|^2\Big) = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\sum_n \mathsf{p}(a_n,t)\Big) = 0$$

La somme des probabilités $\mathsf{p}(a_n, t)$ doit toujours être égale à un. La conservation de la norme de $|\varphi\rangle$ implique, dans un espace de Hilbert de dimension finie, que l'évolution $|\varphi(t = 0)\rangle \to |\varphi(t)\rangle$ est unitaire : on parle alors *d'évolution hamiltonienne*, ou *d'évolution unitaire*. Ce résultat reste vrai en dimension infinie.

La forme matricielle de l'équation d'évolution (4.11) s'obtient dans une base arbitraire $\{|\alpha\rangle\}$ de $\mathcal{H}$ en la multipliant à gauche (4.11) par $\langle\alpha|$ et en utilisant la relation de fermeture

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,\langle\alpha|\varphi(t)\rangle = \langle\alpha|H(t)|\varphi(t)\rangle = \sum_\beta\langle\alpha|H(t)|\beta\rangle\langle\beta|\varphi(t)\rangle$$

soit

$$\mathrm{i}\hbar\dot{c}_\alpha(t) = \sum_\beta H_{\alpha\beta}(t)\,c_\beta(t) \qquad (4.13)$$

Nous avons souligné le caractère réversible et unitaire de l'évolution (4.11). Ce caractère réversible et unitaire doit être contrasté avec celui de l'évolution

(4.7) dans une mesure, qui est *non unitaire et irréversible*, parce que l'évolution limitée à l'espace des états du système pendant la mesure est une évolution tronquée qui se focalise sur l'évolution du système, lequel ne reste pas fermé pendant la mesure. L'évolution globale comprend en outre celle de l'appareil de mesure et de l'environnement (section 11.4). Lorsque l'on examine l'évolution d'un système ouvert, on rencontre automatiquement une évolution non unitaire, qui sera traitée en détail au chapitre 18. Une évolution non unitaire peut parfois être représentée par un hamiltonien non hermitien, par exemple dans la description de particules instables : *cf.* l'exercice 4.4.10.

4.2.2 Opérateur d'évolution

Nous avons donné en (4.11) l'équation d'évolution sous forme différentielle. Il existe une formulation intégrale de cette équation qui fait intervenir l'opérateur d'évolution $U(t, t_0)$. Dans cette formulation, le postulat **IV** devient :

Postulat IV' : Opérateur d'évolution
Le vecteur d'état $|\varphi(t)\rangle$ au temps t se déduit du vecteur d'état $|\varphi(t_0)\rangle$ au temps t_0 par application d'un opérateur unitaire $U(t, t_0)$, appelé *opérateur d'évolution*

$$\boxed{|\varphi(t)\rangle = U(t, t_0)|\varphi(t_0)\rangle} \tag{4.14}$$

L'unitarité de U : $U^\dagger U = UU^\dagger = I$ assure la conservation (4.12) de la norme

$$\langle\varphi(t)|\varphi(t)\rangle = \langle\varphi(t_0)|U^\dagger(t, t_0)U(t, t_0)|\varphi(t_0)\rangle = \langle\varphi(t_0)|\varphi(t_0)\rangle = 1$$

Inversement on aurait pu partir de la conservation de la norme pour montrer que $U^\dagger U = I$. Dans un espace vectoriel de dimension finie, cela suffit à assurer $UU^\dagger = I$ (*cf.* § 2.2.1), mais pas nécessairement dans un espace de dimension infinie. L'opérateur d'évolution obéit aussi à la *propriété de groupe*

$$U(t, t_1)U(t_1, t_0) = U(t, t_0) \qquad t_0 \leq t_1 \leq t \tag{4.15}$$

En effet, il est équivalent d'aller directement de t_0 à t, ou d'aller d'abord de t_0 à t_1 et ensuite de t_1 à t

$$\begin{aligned} |\varphi(t)\rangle &= U(t, t_0)|\varphi(t_0)\rangle \\ &= U(t, t_1)|\varphi(t_1)\rangle = U(t, t_1)U(t_1, t_0)|\varphi(t_0)\rangle \end{aligned}$$

Comme précédemment la restriction $t_0 < t_1 < t$ n'est pas nécessaire : t_1 peut être quelconque. Évidemment $U(t_0, t_0) = I$, et la propriété de groupe jointe à l'unitarité de U implique

$$U(t, t_0) = U^{-1}(t_0, t) = U^\dagger(t_0, t) \tag{4.16}$$

Les postulats d'évolution temporelle **IV** et **IV'** ne sont bien sûr pas indépendants. En effet, il est facile à partir de (4.11) d'écrire une équation différentielle

pour $U(t,t_0)$. En différentiant (4.14) par rapport au temps

$$i\hbar \frac{d}{dt} |\varphi(t)\rangle = i\hbar \left[\frac{d}{dt} U(t,t_0)\right] |\varphi(t_0)\rangle$$

et en comparant avec (4.11), on obtient

$$i\hbar \left[\frac{d}{dt} U(t,t_0)\right] |\varphi(t_0)\rangle = H(t)U(t,t_0)|\varphi(t_0)\rangle$$

Comme cette équation doit être valable quel que soit $|\varphi(t_0)\rangle$, on en déduit une équation différentielle pour $U(t,t_0)$

$$\boxed{i\hbar \frac{d}{dt} U(t,t_0) = H(t)U(t,t_0)} \tag{4.17}$$

ce qui se traduit aussi par

$$H(t_0) = i\hbar \frac{d}{dt} U(t,t_0)\Big|_{t=t_0} \tag{4.18}$$

en prenant la limite $t \to t_0$. Il est donc aisé de passer de la formulation intégrale (4.14) à la formulation différentielle (4.11). Le passage inverse est plus compliqué : en effet, si $H(t)$ était un nombre, l'équation (4.17) s'intègrerait immédiatement. Mais $H(t)$ est un opérateur et en général

$$U(t,t_0) \neq \exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} H(t')\,dt'\right) \tag{4.19}$$

parce qu'il n'y a aucune raison pour que $[H(t'), H(t'')] = 0$. Cependant il existe une formule générale[13] pour calculer $U(t,t_0)$ à partir de $H(t)$, et les postulats **IV** et **IV'** sont strictement équivalents[14].

Revenons maintenant sur la nécessité d'introduire un postulat supplémentaire **IV** ou **IV'** pour l'évolution temporelle (Scarani [2011]). À première vue, il semble que l'on pourrait se fonder sur le théorème de Wigner (chapitre 7) pour déduire l'existence d'un opérateur unitaire effectuant les translations de temps. Cependant, Peres [1993], section 8.6, montre qu'il n'y a aucune raison d'exiger qu'une translation de temps conserve le module du produit scalaire et soit représentée par un opérateur unitaire. Peres donne également un exemple de physique classique pour mettre en évidence cette difficulté. Si l'on veut utiliser le théorème de Wigner, il faut donc une hypothèse supplémentaire, et l'hypothèse la plus faible semble être celle d'une évolution réversible, ce qui veut dire que si une évolution $t_0 \to t > t_0$ est physiquement possible, alors l'évolution $t \to t_0$ l'est également. L'exemple de Peres ne vérifie pas cette

13. Voir par exemple Messiah [1959], chapitre XVII.

14. En toute rigueur, on peut trouver des exceptions où U est défini, mais non H : *cf.* Peres [1993], page 85.

propriété de réversibilité. Cette inversion du sens du temps est examinée en détail dans l'appendice A2. Pour exploiter l'hypothèse de réversibilité, nous allons revenir à la notion de "discernabilité" de deux états quantiques. Supposons qu'un système quantique puisse exister dans deux états possibles, $|\varphi\rangle$ et $|\chi\rangle$. Le module du produit scalaire $S = |\langle\chi|\varphi\rangle|$ de ces deux états quantiques caractérise leur discernabilité : si $S = 0$, $|\varphi\rangle$ et $|\chi\rangle$ sont orthogonaux et la mesure du projecteur $\mathcal{P} = |\varphi\rangle\langle\varphi|$, par exemple, a pour résultat $\mathcal{P} = 1$ si le système est dans l'état $|\varphi\rangle$ et $\mathcal{P} = 0$ s'il est dans l'état $|\chi\rangle$: les deux états sont parfaitement discernables dans une mesure. Inversement, si $S = 1$, les deux états sont identiques. En général, nous avons vu en (4.2) que S est bien une mesure quantitative de la discernabilité des deux états. Soit S_0 la valeur de S au temps $t = t_0$, et supposons que S diminue pour $t > t_0$: $S < S_0$. Dans ce cas, en raison de (4.2), une mesure faite à $t = t_0$ ne donnerait pas la discernabilité optimale : il suffirait d'attendre pour obtenir une meilleure discernabilité que celle obtenue à t_0. La conclusion est que S ne peut pas diminuer entre t_0 et t, S peut seulement augmenter. Mais d'après l'hypothèse d'évolution réversible, si une évolution $t_0 \to t$ est possible, alors l'évolution $t \to t_0$ doit également être physiquement possible. Si S croît entre t et t_0, ce qui est la seule possibilité, alors dans l'évolution $t \to t_0$ S décroît, ce qui est impossible. La seule solution est que S soit indépendant du temps

$$|\langle\chi(t)|\varphi(t)\rangle| = |\langle\chi(t_0)|\varphi(t_0)\rangle|$$

D'après le théorème de Wigner (chapitre 7 et Appendice A1), cela implique que $|\varphi(t)\rangle$ se déduise de $|\varphi(t_0)\rangle$ par une transformation unitaire ou anti-unitaire. Comme la transformation identité pour $t = t_0$ fait partie des transformations admissibles, seule une transformation unitaire $U(t, t_0)$ est possible

$$|\varphi(t)\rangle = U(t, t_0)|\varphi(t_0)\rangle$$

ce qui est le postulat d'évolution sous la forme **IV'**.

4.2.3 États stationnaires

Un cas particulier très important est celui du système isolé. L'opérateur d'évolution ne peut alors pas dépendre du choix fait pour l'origine des temps : peu importe, pour un système isolé de toute influence extérieure, que nous choisissions pour le décrire le temps de Paris ou celui de New-York qui, comme chacun sait, sont décalés de $\tau =$ six heures

$$t_{\rm New-York} = t_{\rm Paris} - \tau$$

Quel que soit τ, nous devons avoir

$$U(t - \tau, t_0 - \tau) = U(t, t_0) \tag{4.20}$$

Ceci implique que U ne peut dépendre que de la *différence* $(t - t_0)$. L'équation (4.18) montre alors que le hamiltonien est *indépendant du temps*, car le

choix de t_0 est arbitraire. Naturellement, il peut parfaitement arriver que le hamiltonien soit indépendant du temps, même pour un système non isolé, par exemple si le système est plongé dans un champ magnétique indépendant du temps comme le spin 1/2 du § 3.2.5. En revanche, si un champ magnétique est appliqué entre 12h et 12h10, heure de Paris, le choix de l'origine des temps ne sera pas indifférent ! Lorsque le hamiltonien est indépendant du temps, l'équation différentielle (4.17) s'intègre sans problème et

$$\boxed{U(t,t_0) = \exp\left(-\frac{\mathrm{i}(t-t_0)}{\hbar}\,H\right)} \tag{4.21}$$

qui ne dépend que de $(t - t_0)$.

L'opérateur $U(t - t_0)$ (4.21) est obtenu par exponentiation de l'opérateur hermitique H ; $U(t - t_0)$ effectue une translation de temps de $(t - t_0)$ sur le vecteur d'état, et si $(t - t_0)$ devient infinitésimal

$$U(t-t_0) \simeq I - \frac{\mathrm{i}(t-t_0)}{\hbar}\,H \tag{4.22}$$

Cette équation s'interprète ainsi : H est le *générateur infinitésimal des translations de temps*, et, pour un sytème isolé, la définition la plus générale du hamiltonien est d'être précisément ce générateur infinitésimal. La notion de générateur infinitésimal sera étendue à d'autres transformations au chapitre 7.

Considérons un système physique isolé qui peut être décrit à une bonne approximation par un vecteur d'état d'un espace de Hilbert de dimension 1 : particule élémentaire stable, atome dans son état fondamental ... Le vecteur d'état est alors un nombre complexe $\varphi(t)$ et H un nombre réel : $H = E$. La loi d'évolution (4.13) devient, compte tenu de (4.20)

$$\varphi(t) = \exp\left(-\frac{\mathrm{i}}{\hbar}E(t-t_0)\right)\varphi(t_0) = \exp(-\mathrm{i}\omega(t-t_0))\varphi(t_0) \tag{4.23}$$

en définissant $E = \hbar\omega$. D'après la relation de Planck-Einstein $E = \hbar\omega$, il est naturel d'identifier E à l'énergie.

Passons maintenant à un cas moins trivial. Soit $|n,r\rangle$ un vecteur propre de H correspondant à la valeur propre E_n : $H|n,r\rangle = E_n|n,r\rangle$. Son évolution temporelle est particulièrement simple : si $|\varphi(t_0)\rangle = |n,r\rangle$

$$|\varphi(t)\rangle = \exp\left(-\frac{\mathrm{i}(t-t_0)}{\hbar}\,H\right)|n,r\rangle = \exp\left(-\frac{\mathrm{i}}{\hbar}\,E_n(t-t_0)\right)|n,r\rangle \tag{4.24}$$

La probabilité de trouver $|\varphi(t)\rangle$ dans un état $|\chi\rangle$ quelconque est indépendante du temps

$$|\langle\chi|\varphi(t)\rangle|^2 = \left|\langle\chi|\exp\left(-\frac{\mathrm{i}}{\hbar}\,E_n(t-t_0)\right)|\varphi(t_0)\rangle\right|^2 = |\langle\chi|\varphi(t_0)\rangle|^2$$

Pour cette raison, un état propre de H est appelé *état stationnaire*.

Il est parfois utile d'écrire la loi d'évolution temporelle sous forme de composantes. Écrivons la décomposition d'un vecteur d'état arbitraire $|\varphi(t_0)\rangle$ au temps $t = t_0$ sur la base $\{|n, r\rangle\}$ des vecteurs propres de H

$$|\varphi(t_0)\rangle = \sum_{n,r} c_{nr}(t_0)|n, r\rangle \qquad c_{nr}(t_0) = \langle n, r|\varphi(t_0)\rangle$$

Nous avons alors

$$\begin{aligned}|\varphi(t)\rangle &= \sum_{n,r} c_{nr}(t_0) \exp\left(-\frac{\mathrm{i}(t-t_0)}{\hbar}\,H\right)|n, r\rangle \\ &= \sum_{n,r} c_{nr}(t_0) \exp\left(-\frac{\mathrm{i}}{\hbar}\,E_n(t-t_0)\right)|n, r\rangle\end{aligned}$$

ce qui donne la variation des coefficients c_{nr} en fonction de t

$$c_{nr}(t) = \exp\left(-\frac{\mathrm{i}}{\hbar}\,E_n(t-t_0)\right) c_{nr}(t_0) \tag{4.25}$$

4.2.4 Inégalité de Heisenberg temporelle

Au § 3.2.5, nous avons donné une explication élémentaire de la relation entre un temps caractéristique d'évolution Δt et une dispersion sur l'énergie ΔE. Établissons maintenant de façon générale une inégalité sur le produit $\Delta E\,\Delta t$, ou *inégalité de Heisenberg temporelle*. Nous allons d'abord écrire l'équation d'évolution de la valeur moyenne $\langle A\rangle_\varphi(t) = \langle\varphi(t)|A|\varphi(t)\rangle$ de l'opérateur A représentant la propriété physique $\mathcal{A}$, supposée indépendante du temps

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\,\langle\varphi(t)|A|\varphi(t)\rangle &= \frac{1}{\mathrm{i}\hbar}\,[-\langle\varphi(t)|HA|\varphi(t)\rangle + \langle\varphi(t)|AH|\varphi(t)\rangle] \\ &= \frac{1}{\mathrm{i}\hbar}\,\langle\varphi(t)|AH - HA|\varphi(t)\rangle\end{aligned}$$

ce qui donne le *théorème d'Ehrenfest*

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\,\langle A\rangle_\varphi(t) = \frac{1}{\mathrm{i}\hbar}\,\langle\varphi(t)|[A, H]|\varphi(t)\rangle = \frac{1}{\mathrm{i}\hbar}\,\langle[A, H]\rangle_\varphi} \tag{4.26}$$

Utilisons maintenant la relation (4.10), en remplaçant B par H

$$\Delta_\varphi H\,\Delta_\varphi A \geq \frac{1}{2}\,|\langle[A, H]\rangle_\varphi| = \frac{1}{2}\,\hbar\left|\frac{\mathrm{d}}{\mathrm{d}t}\,\langle A\rangle_\varphi(t)\right|$$

et définissons le temps $\tau_\varphi(A)$ par

$$\frac{1}{\tau_\varphi(A)} = \left|\frac{\mathrm{d}\langle A\rangle_\varphi(t)}{\mathrm{d}t}\right|\frac{1}{\Delta_\varphi A}$$

$\tau_\varphi(A)$ est le temps caractéristique nécessaire pour que la valeur moyenne de A varie de $\Delta_\varphi A$, c'est-à-dire d'une quantité de l'ordre de la dispersion. L'inégalité précédente devient

$$\Delta_\varphi H\, \tau_\varphi(A) \geq \frac{1}{2}\,\hbar \tag{4.27}$$

ce qui est la forme rigoureuse de l'inégalité de Heisenberg temporelle. Cette inégalité est souvent écrite sous la forme

$$\Delta E\, \Delta t \gtrsim \frac{1}{2}\,\hbar \tag{4.28}$$

ΔE représente la dispersion en énergie et Δt un temps caractéristique d'évolution[15]. La valeur de l'énergie ne peut être exactement fixée que si la dispersion ΔE est nulle, ce qui implique que le temps caractéristique doit être infini. Ceci n'est possible que si l'état du système est un état stationnaire. C'est le cas par exemple pour une particule élémentaire stable ou un atome dans son état fondamental, en l'absence de perturbations extérieures. En revanche, un atome porté dans un état excité n'est pas dans un état stationnaire. En raison de son couplage avec les fluctuations du vide du champ électromagnétique (*cf.* le § 17.3.4), il émet un photon au bout d'un temps moyen τ, appelé *vie moyenne de l'état excité* (*cf.* le § 1.5.3). L'énergie du photon final présente une dispersion en énergie ΔE, qui est appelée *largeur de raie* et est souvent notée $\hbar\Gamma$. Nous verrons ci-dessous que ΔE (ou Γ) et τ sont reliés par une relation qui rappelle l'inégalité de Heisenberg temporelle (4.28), mais qui, en fait, est d'origine *différente*

$$\boxed{\tau\Delta E \simeq \hbar \quad \text{ou} \quad \tau\Gamma \simeq 1} \tag{4.29}$$

Donnons un ordre de grandeur typique de vie moyenne en physique atomique, en prenant comme exemple le premier niveau excité de l'atome de rubidium ; l'atome dans cet état excité revient à son état fondamental en émettant un photon de longueur d'onde $\lambda = 0.78\,\mu$m, ce qui correspond à une énergie $\varepsilon = 1.6\,$eV. La largeur de raie est $\hbar\Gamma = 2.4 \times 10^{-8}\,$eV, et la vie moyenne $\tau \simeq 1/\Gamma = 2.7 \times 10^{-8}\,$s. La dispersion sur l'énergie de l'état excité est donc très faible par rapport à la différence d'énergie avec le niveau fondamental : $\hbar\Gamma/\varepsilon \simeq 10^{-8}$, ce qui entraîne que l'énergie du niveau excité est définie avec une précision excellente. La relation (4.29) se généralise à toute désintégration de particules, par exemple la désintégration à deux corps $C \to A + B$.

Il ne faudrait surtout pas conclure de (4.28) que l'on ne peut pas mesurer l'énergie avec une précision meilleure que ΔE. Considérons par exemple l'énergie E du boson Z^0, vecteur de l'interaction faible (*cf.* § 1.1.4) dans son

15. L'inégalité $\Delta E\, \Delta t \gtrsim \hbar$ a un statut différent de (4.10) dans la mesure où t n'est pas un opérateur (exercice 4.4.5). Δt est souvent interprété incorrectement comme le temps nécessaire à la mesure de l'énergie.

référentiel au repos : $E = m_Z c^2$, où m_Z est la masse du boson Z^0. Le boson Z^0 est instable, et possède donc une largeur de raie. Celle-ci a été mesurée avec une grande précision : $\hbar\Gamma_Z = 2.4952 \pm 0.0023\,\text{GeV}$. La précision actuelle sur la masse du Z^0 est bien meilleure que Γ_Z ! En effet, la détermination la plus précise est actuellement $m_Z c^2 = 91.1875 \pm 0.0021\text{GeV}$ (figure 4.3). En d'autres termes, il est possible de pointer le centre de la raie avec une précision bien meilleure que la dispersion.

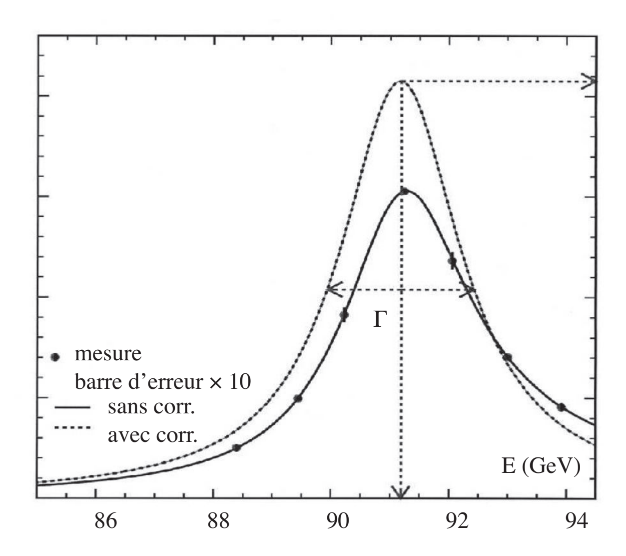


FIG. 4.3 – Spectre de masse du boson Z^0. La courbe en trait plein est le résultat expérimental brut. Ce résultat doit être corrigé pour tenir compte des corrections radiatives (émission de photons) exactement calculables. La courbe en pointillés donne le spectre de masse du Z^0. D'après la collaboration LEP, prétirage CERN EP-2000-13 (2000).

Revenons maintenant à la discussion de (4.29). Soit $|\varphi\rangle$ le vecteur d'état à $t = 0$ d'un état quantique instable, par exemple l'état excité d'un atome ou l'état d'une particule instable. Pour fixer les idées, prenons le cas du niveau excité d'un atome. Cet atome est supposé isolé d'influences extérieures, mais il ne peut pas être isolé du champ électromagnétique quantifié. Le hamiltonien H qui décrit son retour au niveau fondamental par émission spontanée d'un photon est indépendant du temps. Au temps $t = 0$, le vecteur d'état du système global est $|\varphi\rangle = |\psi(t = 0)\rangle$ (il n'y a pas de photon), et il devient $|\psi(t)\rangle$ au temps t : $|\psi(t)\rangle$ décrit l'évolution de l'ensemble atome + champ, et cette évolution obéit à (4.11) dans l'espace des états atome + champ

$$|\psi(t)\rangle = \exp\left(-\frac{\mathrm{i}Ht}{\hbar}\right)|\varphi\rangle \tag{4.30}$$

Soit $|n\rangle$ une base d'états propres de H, $H|n\rangle = E_n|n\rangle$; on note que $|\varphi\rangle$ n'est pas un tel état propre, car il n'est pas stationnaire. La probabilité $\mathsf{p}(t)$ de trouver l'atome au temps t dans son état excité, ou *probabilité de survie*, est

$$\mathsf{p}(t) = |c(t)|^2 = |\langle\varphi|\psi(t)\rangle|^2 = |\langle\varphi|\, \mathrm{e}^{-\mathrm{i}Ht/\hbar}\, |\varphi\rangle|^2 \tag{4.31}$$

Il est possible de montrer (exercice 4.4.5) l'inégalité suivante

$$\mathsf{p}(t) \geq \cos^2\left(\frac{t\Delta H}{\hbar}\right) \qquad 0 \leq t \leq \frac{\pi\hbar}{2\Delta H} \tag{4.32}$$

où $\Delta H = (\langle H^2\rangle - \langle H\rangle^2)^{1/2}$ est la dispersion de H *calculée dans l'état* $|\varphi\rangle$. Pour des raisons que nous verrons dans un instant, cette inégalité n'est pas très utile en pratique. En insérant dans (4.31) la relation de fermeture $\sum |n\rangle\langle n| = I$, on obtient pour l'amplitude de probabilité $c(t)$

$$c(t) = \sum_n |\langle\varphi|n\rangle|^2 \mathrm{e}^{-\mathrm{i}E_n t/\hbar}$$

Soit $w(E)$ la tranformée de Fourier de $c(t)$, ou fonction spectrale

$$\begin{aligned} w(E) &= \int \frac{\mathrm{d}t}{2\pi}\, \mathrm{e}^{\mathrm{i}Et/\hbar}\, c(t) = \sum_n \int \frac{\mathrm{d}t}{2\pi}\, \mathrm{e}^{\mathrm{i}(E-E_n)t/\hbar} |\langle\varphi|n\rangle|^2 \\ &= \hbar \sum_n |\langle\varphi|n\rangle|^2 \delta(E - E_n) \end{aligned} \tag{4.33}$$

On remarque que la fonction spectrale $w(E)$ est positive : $w(E) \geq 0$. Il est immédiat de calculer les valeurs moyennes de H et de H^2

$$\begin{aligned} \langle H\rangle &= \sum_n E_n\, |\langle\varphi|n\rangle|^2 = \int \mathrm{d}E\, E\, w(E) \\ \langle H^2\rangle &= \sum_n E_n^2\, |\langle\varphi|n\rangle|^2 = \int \mathrm{d}E\, E^2\, w(E) \end{aligned} \tag{4.34}$$

L'approximation de Wigner et Weisskopf (annexe B) montre qu'une forme approchée de $w(E)$ est

$$w(E) = \frac{\Gamma\hbar}{2\pi} \frac{1}{(E - E_0)^2 + \hbar^2\Gamma^2/4} \tag{4.35}$$

de sorte que, par transformée de Fourier inverse,

$$c(t) = \mathrm{e}^{-\mathrm{i}E_0 t/\hbar}\, \mathrm{e}^{-\Gamma t/2} \tag{4.36}$$

et la probabilité de survie

$$\mathsf{p}(t) = |c(t)|^2 \mathrm{e}^{-\Gamma t} \tag{4.37}$$

est (approximativement) une exponentielle. Insistons sur le fait que la forme (4.35) est seulement approchée, et que les lois (4.36) et (4.37) ne peuvent pas être exactes. L'inégalité (4.32) montre qu'elle ne peuvent pas être valables pour des temps trop courts, et on s'attend aussi à des déviations pour des temps très longs. L'argument développé ci-dessus montre aussi que la largeur de raie ΔE n'est en rien liée à la dispersion ΔH. En fait, la forme lorentzienne (4.35) donne une dispersion infinie, ce qui est difficilement acceptable car cela voudrait dire que $|\varphi\rangle$ n'appartient pas au domaine de H, et l'équation d'évolution (4.11) ne pourrait pas exister[16]. En fait, on peut montrer que la loi exponentielle est approximativement valable[17] pourvu que $\Delta E \ll \Delta H$.

La relation (4.28) permet aussi de discuter la notion de "particules virtuelles". Il est possible d'interpréter les processus de la théorie quantique des champs en termes d'échanges de particules virtuelles : par exemple l'interaction coulombienne dans l'atome d'hydrogène correspond à l'échange de photons virtuels entre un proton et un électron. Ces processus ne correspondent pas à une réaction observable entre particules, car les particules virtuelles ne peuvent pas obéir à la fois à la conservation de l'énergie-impulsion et à la condition reliant énergie-impulsion et masse : $E^2 = \vec{p}^2 c^2 + m^2 c^4$. Prenons l'exemple des interactions entre nucléons, ou interactions fortes (*cf.* le § 1.1.4), dont Yukawa imagina vers 1935 qu'elles étaient dues à l'échange d'une particule encore inconnue à l'époque, et que nous appelons aujourd'hui méson π. Cet échange est représenté sur la figure 4.4 par un "diagramme de Feynman". Le proton de gauche (p) émet un méson π^+ et se transforme en neutron (n), tandis que le neutron de droite absorbe ce méson π^+ et se transforme en proton. La conservation de l'énergie-impulsion interdit la réaction

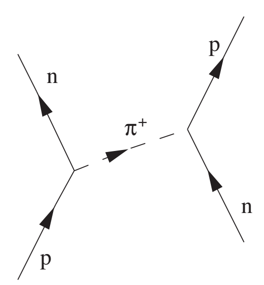


FIG. 4.4 – Diagramme de Feynman pour l'échange d'un méson π.

$$p \to n + \pi^+$$

Si l'impulsion est conservée, alors l'énergie ne peut pas l'être. En revanche, si l'on admet que la réaction ne dure qu'un temps très court Δt, alors il est possible de tirer parti d'une fluctuation d'énergie $\Delta E \simeq \hbar/\Delta t$. La fluctuation

16. Voir la note 14.

17. Les conditions de validité de la loi exponentielle sont examinées par Peres [1980].

d'énergie nécessaire pour que la réaction soit possible est $\Delta E \sim m_\pi c^2$, où m_π est la masse du méson π^+. Dans l'intervalle de temps Δt, le méson peut parcourir au maximum une distance[18] $\sim c\Delta t \sim \hbar/(m_\pi c)$, la longueur d'onde Compton du méson π. Cette distance représente la portée maximale r_0 des forces nucléaires (*cf.* § 1.1.4), qui est de l'ordre de 1 fm. Yukawa fut donc capable de prédire l'existence d'une particule ayant une masse de l'ordre de $\hbar/(cr_0) \sim 200$ MeV, et le méson π fut effectivement découvert quelques années plus tard avec une masse de 140 MeV. Le méson π échangé sur la figure 4.4 n'est pas observable : il est virtuel. On sait aujourd'hui que les forces nucléaires ne sont pas fondamentales, et qu'elles sont dérivées de forces fondamentales entre quarks. L'argument de Yukawa reste néanmoins valable, car on peut écrire une théorie effective des forces nucléaires avec échange de mésons, et leur portée maximale est déterminée par le méson le plus léger, qui est le méson π. Le photon étant de masse nulle, la portée des forces électromagnétiques est infinie : comme nous l'avons noté au § 1.1.4, le potentiel coulombien est à longue portée.

4.2.5 Points de vue de Schrödinger et de Heisenberg

Le point de vue adopté dans ce qui précède, où le vecteur d'état évolue avec le temps alors que les opérateurs, sauf dans le cas d'une interaction avec un champ extérieur variable, sont indépendants du temps, est appelé *point de vue de Schrödinger*. Un point de vue équivalent en ce qui concerne les résultats physiques est celui de Heisenberg, où les vecteurs d'état sont indépendants du temps et les opérateurs dépendent du temps. Afin de simplifier la discussion, nous prenons le cas d'un hamiltonien H et d'un opérateur A indépendants du temps. Ce n'est pas le cas général, car il peut arriver que même dans le point de vue de Schrödinger un opérateur A ait une dépendance explicite par rapport au temps, ou que H dépende du temps. Nous admettrons que tel n'est pas le cas, en renvoyant l'étude générale à l'exercice 4.4.7. La valeur moyenne de A au temps t est

$$\langle A\rangle_\varphi(t) = \langle\varphi(t_0)|\exp\left(\frac{\mathrm{i}(t-t_0)}{\hbar}H\right)A\exp\left(-\frac{\mathrm{i}(t-t_0)}{\hbar}H\right)|\varphi(t_0)\rangle$$

Si nous définissons *l'opérateur A dans le point de vue de Heisenberg* $A_{\mathrm{H}}(t)$ par

$$\boxed{A_{\mathrm{H}}(t) = \exp\left(\frac{\mathrm{i}(t-t_0)}{\hbar}H\right)A\exp\left(-\frac{\mathrm{i}(t-t_0)}{\hbar}H\right) \qquad A(t_0)=A} \tag{4.38}$$

alors la valeur moyenne de A peut se calculer comme

$$\langle A\rangle_\varphi(t) = \langle\varphi(t_0)|A_{\mathrm{H}}(t)|\varphi(t_0)\rangle \tag{4.39}$$

18. On néglige pour simplifier la dilatation du temps.

La dépendance par rapport au temps est intégrée dans l'opérateur, tandis que le vecteur d'état est indépendant de t.

Un point de vue intermédiaire entre celui de Schrödinger et celui de Heisenberg est le *point de vue de l'interaction* (ou de Dirac). Ce point de vue est utile lorsqu'il est naturel de décomposer le hamiltonien H en un “hamiltonien libre” H_0 indépendant du temps, que l'on sait diagonaliser, et un hamiltonien d'interaction $W(t)$. L'objectif est de se débarrasser de l'évolution connue de H_0. On définit le vecteur d'état $|\tilde{\varphi}(t)\rangle$ dans le point de vue de l'interaction par

$$|\tilde{\varphi}(t)\rangle = \mathrm{e}^{\mathrm{i}H_0 t/\hbar}\,|\varphi(t)\rangle \qquad |\tilde{\varphi}(t=0)\rangle = |\varphi(t=0)\rangle \tag{4.40}$$

Nous avons choisi $t_0 = 0$ afin d'alléger les notations ; le point de vue de l'interaction coïncide avec celui de Heisenberg si $W = 0$. L'opérateur d'évolution $U(t) = U(t,0)$ vérifie

$$\mathrm{i}\hbar\frac{\mathrm{d}U}{\mathrm{d}t} = [H_0 + W(t)]U(t) \tag{4.41}$$

et on écrit

$$U(t) = U_0(t)\,\tilde{U}(t) \qquad U_0(t) = \mathrm{e}^{-\mathrm{i}H_0 t/\hbar} \tag{4.42}$$

$\tilde{U}(t)$ est l'opérateur d'évolution dans le point de vue de l'interaction. L'équation (4.41) devient

$$\mathrm{i}\hbar\left[\frac{\mathrm{d}U_0}{\mathrm{d}t}\,\tilde{U} + U_0\frac{\mathrm{d}\tilde{U}}{\mathrm{d}t}\right] = [H_0 + W(t)]U_0(t)\tilde{U}(t) \tag{4.43}$$

Les deux premiers termes de chaque membre de (4.43) se compensent, et il reste

$$\begin{aligned}\mathrm{i}\hbar\frac{\mathrm{d}\tilde{U}}{\mathrm{d}t} &= \left[\mathrm{e}^{\mathrm{i}H_0 t/\hbar}\,W(t)\,\mathrm{e}^{-\mathrm{i}H_0 t/\hbar}\right]\tilde{U}(t) = \tilde{W}(t)\tilde{U}(t)\\ \tilde{W}(t) &= \mathrm{e}^{\mathrm{i}H_0 t/\hbar}\,W(t)\,\mathrm{e}^{-\mathrm{i}H_0 t/\hbar}\end{aligned} \tag{4.44}$$

Les équations (4.44) donnent l'évolution dans le point de vue de l'interaction.

4.3 Approximations et modélisation

Nous avons énoncé ci-dessus les principes généraux qui fixent le cadre universel de la théorie quantique. Cela ne veut pas dire que nous sommes prêts à aborder immédiatement un problème physique ! En effet, pour aborder un problème concret, par exemple celui du calcul des niveaux d'énergie de l'atome d'hydrogène, nous avons besoin de fixer l'espace des états et le hamiltonien appropriés selon le degré de précision avec lequel nous souhaitons résoudre le problème. Le choix d'un espace des états et d'un hamiltonien implique toujours que l'on se place dans le cadre d'une certaine approximation, et il ne faut surtout pas confondre ce qui est approximation (ou modélisation) et ce

qui est principe fondamental. Par exemple, ainsi que nous allons le montrer dans un instant, l'espace des états est toujours au départ de dimension infinie, mais il peut arriver qu'il soit possible de se placer de façon approchée dans un espace des états de dimension finie, qui peut même éventuellement être petite ; la dimension N de cet espace est appelée le *nombre de niveaux* de l'approximation. Nous en avons vu un exemple dans l'étude du spin 1/2 : en première approximation, les degrés de liberté de spin sont découplés des degrés de liberté d'espace, et c'est ce qui nous a permis de nous placer dans un espace à deux dimensions en ignorant les degrés de liberté spatiaux. Un autre exemple est celui de l'atome à deux niveaux, modèle standard de la physique atomique : lorsque l'on s'intéresse à l'interaction d'un atome avec un champ électromagnétique de fréquence ω, en pratique le champ d'un laser, et si deux niveaux d'énergie sont espacés de $\hbar\omega_0 \simeq \hbar\omega$, on peut se restreindre à ces deux niveaux d'énergie formant une base pour un espace des états à deux dimensions, et écrire un hamiltonien approché d'interaction avec le champ laser agissant dans cet espace : *cf.* les § 5.4 et § 15.3.1. Cette approche fournit une excellente approximation pour l'interaction laser-atome, approche qui peut être facilement raffinée, par exemple s'il faut tenir compte des effets du spin des niveaux.

Malheureusement, la situation n'est pas toujours aussi simple. Nous allons le voir dans le cas des degrés de liberté spatiaux, que l'on peut traiter en s'appuyant sur le *principe de correspondance.* Selon ce principe, les propriétés physiques position et impulsion sont des opérateurs hermitiens $\vec{R}$ et $\vec{P}$, de composantes X_i et P_j, $i, j = (x, y, z)$, qui vérifient les relations de commutation dites *relations de commutation canoniques*

$$[X_i, P_j] = \mathrm{i}\hbar\delta_{ij}I \tag{4.45}$$

Prenant la trace des deux membres, on observe qu'il est impossible que ces relations soient satisfaites dans un espace de dimension finie : en effet la trace du membre de gauche est nulle (la trace d'un commutateur est nulle), tandis que celle du membre de droite est $\mathrm{i}\hbar N$, où N est la dimension de $\mathcal{H}$. Une fois cette difficulté identifiée, la suite de la procédure (qui n'est pas toujours exempte d'ambiguïtés) consiste à remplacer dans l'expression classique de l'énergie E les positions et les impulsions classiques $\vec{r}$ et $\vec{p}$ par les opérateurs $\vec{R}$ et $\vec{P}$ pour obtenir le hamiltonien quantique d'une particule de masse m dont l'énergie potentielle est $V(\vec{r})$. Le principe de correspondance conduit au passage suivant $E \to H$

$$E = \frac{\vec{p}^2}{2m} + V(\vec{r}) \to H = \frac{\vec{P}^2}{2m} + V(\vec{R}) \tag{4.46}$$

Dans le cas de l'atome d'hydrogène, (4.46) fournit une très bonne approximation si l'on prend pour $V(\vec{r})$ le potentiel coulombien (1.3) et pour espace des états celui de l'électron. L'effet de la masse finie du proton est pris en compte

grâce à la masse réduite. Il faut bien comprendre que (4.45) et (4.46) représentent un choix pour l'espace des états et le hamiltonien, et que des approximations ont été faites. On a négligé en particulier les effets relativistes, et les choses se compliquent dès que l'on en tient compte. Il est à la rigueur possible dans un premier temps de généraliser l'expression du hamiltonien (équation de Dirac : chapitre 19), mais une véritable théorie quantique et relativiste implique que l'on introduise un champ électron-positron et un champ électromagnétique quantifiés : c'est ce que l'on appelle *l'électrodynamique quantique.* Dans ces conditions, le principe de correspondance sous la forme (4.45) n'est plus valable[19] : en fait, il n'y a même plus d'opérateur position ! Et l'électrodynamique quantique n'est elle-même qu'une approximation d'une théorie plus vaste... Il faut donc soigneusement distinguer les principes fondamentaux des approximations nécessaires pour aborder tout problème physique concret. Comme le souligne Isham [1995], la procédure standard qui consite à “quantifier une théorie classique” en utilisant le principe de correspondance n'a qu'une valeur heuristique, et en fin de compte les approximations reposant sur ce principe ou toute autre démarche heuristique doivent être validées par la confrontation aux résultats expérimentaux.

Pour conclure, donnons la forme générale du principe de correspondance : en mécanique analytique, on définit pour un système à N degrés de liberté un hamiltonien $H_{\rm cl}(q_i, p_i)$, supposé indépendant du temps pour simplifier, $i = 1, 2, \ldots, N$. Les équations de Hamilton sont

$$\frac{\partial H_{\rm cl}}{{\rm d}q_i} = -\dot{p}_i \qquad \frac{\partial H_{\rm cl}}{{\rm d}p_i} = \dot{q}_i \tag{4.47}$$

Les variables q_i et p_i sont dites *canoniquement conjuguées.* Le principe de correspondance consiste à associer aux variables classiques q_i et p_i des opérateurs Q_i et P_i

$$q_i \to Q_i \qquad p_i \to P_i$$

obéissant aux *relations de commutation canoniques* (RCC)

$$[Q_i, P_j] = {\rm i}\hbar\delta_{ij} I \tag{4.48}$$

Cependant, cette procédure n'est pas exempte d'ambiguïtés, car l'ordre des variables classiques (q_i, p_i) est indifférent, mais pas celui des opérateurs (Q_i, P_i).

Nous avons, jusqu'à présent, utilisé des notations différentes pour une propriété physique $\mathcal{A}$ et l'opérateur hermitien associé A. Nous abandonnons désormais cette distinction, et confondons les notations pour la propriété physique et l'opérateur correspondant, qui seront représentés tous deux – sauf

19. Il est remplacé par des relations de commutation canoniques entre les champs et leurs moments conjugués, ce qui conduit à des objets mathématiques complexes, les distributions à valeurs opérateur. Toutefois, il reste encore un tel chemin à parcourir (invariance de jauge, renormalisation) avant de calculer une quantité physique que ce principe de correspondance apparaît un peu accessoire, et il est d'ailleurs remplacé en pratique par l'approche des intégrales de chemin de Feynman (§ 12.2.3).

mention explicite du contraire – par des lettres majuscules : hamiltonien H, position $\vec{R}$, impulsion $\vec{P}$, moment angulaire $\vec{J}\ldots$ Les valeurs propres seront désignées par les lettres minuscules correspondantes : $\vec{r}$, $\vec{p}$, $\vec{\jmath}\ldots$, à l'exception du cas de l'énergie, où les valeurs propres de H seront notées E.

4.4 Exercices

4.4.1 Dispersion et vecteurs propres

Montrer qu'une condition nécessaire et suffisante pour que $|\varphi\rangle$ soit vecteur propre d'un opérateur hermitique A est que la dispersion (4.8) $\Delta_\varphi A = 0$.

4.4.2 Méthode variationnelle

1. Soit $|\varphi\rangle$ un vecteur (non normalisé) de l'espace de Hilbert des états et un hamiltonien H. La valeur moyenne $\langle H\rangle_\varphi$ est

$$\langle H\rangle_\varphi = \frac{\langle\varphi|H|\varphi\rangle}{\langle\varphi|\varphi\rangle}$$

Montrer que si le minimum de cette valeur moyenne est obtenu pour $|\varphi\rangle = |\varphi_m\rangle$ et le maximum pour $|\varphi\rangle = |\varphi_M\rangle$, alors

$$H|\varphi_m\rangle = E_m|\varphi_m\rangle \quad \text{et} \quad H|\varphi_M\rangle = E_M|\varphi_M\rangle$$

où E_m et E_M sont la plus petite et la plus grande valeur propre.

2. On suppose que le vecteur $|\varphi\rangle$ dépend d'un paramètre α : $|\varphi\rangle = |\varphi(\alpha)\rangle$. Montrer que si

$$\left.\frac{\partial\langle H\rangle_{\varphi(\alpha)}}{\partial\alpha}\right|_{\alpha=\alpha_0} = 0$$

alors $E_m \le \langle H\rangle_{\varphi(\alpha_0)}$ si α_0 correspond à un minimum de $\langle H\rangle_{\varphi(\alpha)}$ et $\langle H\rangle_{\varphi(\alpha_0)} \le E_M$ si α_0 correspond à un maximum. Ce résultat est à la base d'une méthode d'approximation appelée méthode variationnelle (§ 15.1.4).

3. Si H agit dans un espace à deux dimensions, sa forme la plus générale est

$$H = \begin{pmatrix} a+c & b \\ b & a-c \end{pmatrix}$$

où b peut toujours être choisi réel. En paramétrant $|\varphi(\alpha)\rangle$ sous la forme

$$|\varphi(\alpha)\rangle = \begin{pmatrix} \cos\alpha/2 \\ \sin\alpha/2 \end{pmatrix}$$

trouver les valeurs de α_0 en cherchant les extrema de $\langle\varphi(\alpha)|H|\varphi(\alpha)\rangle$. Retrouver ainsi (2.35).

4.4.3 Théorème de Feynman-Hellmann

Soit un hamiltonien H dépendant d'un paramètre λ : $H = H(\lambda)$, $E(\lambda)$ une valeur propre simple et $|\varphi(\lambda)\rangle$ le vecteur propre normalisé ($||\varphi(\lambda)||^2 = 1$) correspondant

$$H(\lambda)|\varphi(\lambda)\rangle = E(\lambda)|\varphi(\lambda)\rangle$$

Montrer le théorème de Feynman-Hellmann

$$\frac{\partial E}{\partial \lambda} = \langle\varphi(\lambda)\Big|\frac{\partial H}{\partial \lambda}\Big|\varphi(\lambda)\rangle \tag{4.49}$$

4.4.4 Évolution temporelle d'un système à deux niveaux

On considère un système à deux niveaux de hamiltonien H représenté par la matrice

$$H = \hbar \begin{pmatrix} A & B \\ B & -A \end{pmatrix}$$

dans la base

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

D'après (2.35), les valeurs propres et vecteurs propres de H sont

$$\begin{aligned} E_+ &= \hbar\sqrt{A^2+B^2} \qquad |\chi_+\rangle = \cos\frac{\theta}{2}\,|+\rangle + \sin\frac{\theta}{2}\,|-\rangle \\ E_- &= -\hbar\sqrt{A^2+B^2} \qquad |\chi_-\rangle = -\sin\frac{\theta}{2}\,|+\rangle + \cos\frac{\theta}{2}\,|-\rangle \end{aligned}$$

avec

$$A = \sqrt{A^2+B^2}\,\cos\theta \qquad B = \sqrt{A^2+B^2}\,\sin\theta \qquad \tan\theta = \frac{B}{A}$$

1. Le vecteur d'état $|\varphi(t)\rangle$ au temps t peut se décomposer sur la base $\{|+\rangle, |-\rangle\}$

$$|\varphi(t)\rangle = c_+(t)|+\rangle + c_-(t)|-\rangle$$

Écrire le système d'équations différentielles couplées auquel obéissent les composantes $c_+(t)$ et $c_-(t)$.

2. On décompose $|\varphi(t=0)\rangle$ sur la base $\{|\chi_+\rangle, |\chi_-\rangle\}$

$$|\varphi(t=0)\rangle = |\varphi(0)\rangle = \lambda|\chi_+\rangle + \mu|\chi_-\rangle \qquad |\lambda|^2 + |\mu|^2 = 1$$

Montrer que $c_+(t) = \langle +|\varphi(t)\rangle$ s'écrit

$$c_+(t) = \lambda\,\mathrm{e}^{-\mathrm{i}\Omega t/2}\cos\frac{\theta}{2} - \mu\,\mathrm{e}^{\mathrm{i}\Omega t/2}\sin\frac{\theta}{2}$$

avec $\Omega = 2\sqrt{A^2+B^2}$: $\hbar\Omega$ est la différence d'énergie entre les deux niveaux. En déduire que $c_+(t)$ (de même que $c_-(t)$) vérifie l'équation différentielle

$$\ddot{c}_+(t) + \left(\frac{\Omega}{2}\right)^2 c_+(t) = 0$$

3. On suppose que $c_+(0) = 0$. En déduire λ et μ à une phase près ainsi que $c_+(t)$. Montrer que la probabilité de trouver le système au temps t dans l'état $|+\rangle$ est

$$\mathsf{p}_+(t) = \sin^2\theta\,\sin^2\left(\frac{\Omega t}{2}\right) = \frac{B^2}{A^2+B^2}\,\sin^2\left(\frac{\Omega t}{2}\right)$$

4. Montrer que si $c_+(t=0) = 1$ alors

$$c_+(t) = \cos\frac{\Omega t}{2} - \mathrm{i}\cos\theta\sin\frac{\Omega t}{2}$$

En déduire $\mathsf{p}_+(t)$ et $\mathsf{p}_-(t)$, et vérifier la compatibilité du résultat avec celui de la question précédente.

4.4.5 Inégalités de Heisenberg temporelles

1. Supposons qu'il existe un opérateur hermitien T obéissant à $[T, H] = \mathrm{i}\hbar I$. En examinant l'état

$$|\Psi\rangle = H\mathrm{e}^{\mathrm{i}\alpha T}|E\rangle$$

où $|E\rangle$ est un état propre de H, $H|E\rangle = E|E\rangle$, montrer que le spectre de H ne serait pas borné inférieurement.

2. Soit $|\varphi\rangle$ le vecteur d'état au temps $t = 0$ d'une particule instable, et soit $\mathsf{p}(t)$ sa probabilité de survie, c'est-à-dire la probabilité que cette particule ne soit pas désintégrée au temps t. La particule est supposée isolée de toute influence extérieure (mais pas des champs quantifiés), de sorte que le hamiltonien H qui régit la désintégration est indépendant du temps. Soit $|\psi(t)\rangle$ le vecteur d'état au temps t de l'ensemble particule + champs quantifiés

$$|\psi(t)\rangle = \exp\left(-\frac{\mathrm{i}Ht}{\hbar}\right)|\varphi\rangle$$

L'amplitude de probabilité pour trouver le système quantique au temps t dans $|\varphi\rangle$ est

$$c(t) = \langle\varphi|\psi(t)\rangle = \left\langle\varphi\right|\exp\left(-\frac{\mathrm{i}Ht}{\hbar}\right)\left|\varphi\right\rangle$$

et la probabilité de survie est

$$\mathsf{p}(t) = |c(t)|^2 = |\langle\psi(t)|\varphi\rangle|^2 = \langle\psi(t)|\mathcal{P}|\psi(t)\rangle\ ,$$

où $\mathcal{P} = |\varphi\rangle\langle\varphi|$ est le projecteur sur l'état initial. On se restreint d'abord à des temps courts. Montrer que pour $t \to 0$

$$\mathsf{p}(t) \simeq 1 - \frac{(\Delta H)^2}{\hbar^2}\, t^2 \ ,$$

de sorte que pour des temps très courts la loi de désintégration n'est certainement pas exponentielle. Les valeurs moyennes de H et H^2 sont calculées avec $|\varphi\rangle$. On remarque que ΔH doit être fini, sinon $|\varphi\rangle$ n'appartiendrait pas au domaine de H^2, ce qui serait dificile à imaginer physiquement (voir le chapitre 6 pour la définition du domaine d'un opérateur).

3. Un résultat plus général est obtenu comme suit. Montrer d'abord que

$$\Delta\mathcal{P}^2 = \langle\mathcal{P}\rangle - \langle\mathcal{P}\rangle^2$$

et utiliser la forme générale de l'inégalité de Heisenberg temporelle (4.27) pour montrer l'inégalité

$$\left|\frac{\mathrm{d}\mathsf{p}(t)}{\mathrm{d}t}\right| \le \frac{2\Delta H}{\hbar}\,\sqrt{\mathsf{p}(1-\mathsf{p})} \ .$$

En intégrant cette équation différentielle, déduire

$$\mathsf{p}(t) \ge \cos^2\left(\frac{t\Delta H}{\hbar}\right) \qquad 0 \le t \le \frac{\pi\hbar}{2\Delta H}$$

4.4.6 L'énigme des neutrinos solaires

Les réactions nucléaires à l'intérieur du Soleil produisent en abondance des neutrinos électroniques ν_e ; 95 % de ces neutrinos sont produits dans la réaction

$$p + p \ \to \ {}^2\mathrm{H} + e^+ + \nu_e$$

La Terre reçoit du Soleil 6.5×10^{14} neutrinos par seconde et par m^2. Depuis une trentaine d'années, plusieurs expériences ont tenté de détecter ces neutrinos, mais toutes ces expériences concordent pour conclure que le flux de neutrinos mesuré est seulement la moitié environ du flux calculé à partir du modèle standard du Soleil. Or, ce modèle est considéré comme particulièrement fiable[20], en particulier en raison des résultats récents de l'héliosismologie : les incertitudes sur le modèle solaire ne peuvent en aucun cas expliquer ce "déficit en neutrinos solaires". La combinaison de trois expériences (*cf.* la note 4 du chapitre 1), a permis de montrer sans aucun doute possible que ce déficit en neutrinos est dû à la transformation des neutrinos ν_e en d'autres sortes de neutrinos au cours de leur voyage entre le Soleil et la Terre : ces expériences montrent que le flux *total* de neutrinos prévu par le modèle du Soleil est correct, mais que c'est le flux de neutrinos électroniques qui est trop faible. Nous allons faire une théorie simplifiée, mais qui donne l'essentiel de la physique sous-jacente, en supposant

20. L'intérieur du Soleil est connu de façon bien plus précise que celui de la Terre !

- qu'il existe seulement deux types de neutrinos, le neutrino électronique ν_e et le neutrino muonique ν_μ (en fait il existe une troisième sorte de neutrino, le neutrino τ : ν_τ)
- que tout le phénomène se passe dans le vide dans la propagation entre le Soleil et la Terre (en fait la propagation dans le Soleil joue aussi un rôle important[21]).

On a longtemps admis que les neutrinos étaient des particules de masse nulle. Si, au contraire, ils sont massifs, on peut se placer dans leur référentiel au repos et écrire le hamiltonien dans la base $\{|\nu_e\rangle, |\nu_\mu\rangle\}$

$$|\nu_e\rangle = \begin{pmatrix}1\\0\end{pmatrix} \qquad |\nu_\mu\rangle = \begin{pmatrix}0\\1\end{pmatrix} \qquad H = c^2 \begin{pmatrix} m_e & m \\ m & m_\mu \end{pmatrix}$$

L'élément non diagonal m permet des transitions entre neutrinos électroniques et neutrinos muoniques.

1. Montrer que les états ayant une masse déterminée sont $|\nu_1\rangle$ et $|\nu_2\rangle$

$$\begin{aligned} |\nu_1\rangle &= \cos\frac{\theta}{2}\,|\nu_e\rangle + \sin\frac{\theta}{2}\,|\nu_\mu\rangle \\ |\nu_2\rangle &= -\sin\frac{\theta}{2}\,|\nu_e\rangle + \cos\frac{\theta}{2}\,|\nu_\mu\rangle \end{aligned}$$

avec

$$\tan\theta = \frac{2m}{m_e - m_\mu}$$

et que les masses m_1 et m_2 sont

$$\begin{aligned} m_1 &= \frac{m_e + m_\mu}{2} + \sqrt{m^2 + \left(\frac{m_e - m_\mu}{2}\right)^2} \\ m_2 &= \frac{m_e + m_\mu}{2} - \sqrt{m^2 + \left(\frac{m_e - m_\mu}{2}\right)^2} \end{aligned}$$

2. Les neutrinos se propagent avec une vitesse proche de celle de la lumière : leur énergie est très grande par rapport à $\langle m\rangle c^2$, où $\langle m\rangle$ est une masse typique figurant dans H. Montrer que si un neutrino électronique est produit au temps $t = 0$ dans le Soleil, le vecteur d'état étant

$$|\varphi(t=0)\rangle = |\nu_e\rangle = \cos\frac{\theta}{2}\,|\nu_1\rangle - \sin\frac{\theta}{2}\,|\nu_2\rangle$$

le vecteur d'état au temps t a pour composante sur $|\nu_e\rangle$

$$\langle \nu_e|\varphi(t)\rangle = \mathrm{e}^{-\mathrm{i}E_1 t/\hbar}\left(\cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2}\,\mathrm{e}^{-\mathrm{i}\Delta E\, t/\hbar}\right)$$

21. Voir Abers [2004], chapitre 6, pour une discussion élémentaire de cette propagation.

où $\Delta E = E_2 - E_1$. En déduire que la probabilité de trouver un neutrino ν_e au temps t est

$$\mathsf{p}_e(t) = 1 - \sin^2\theta \sin^2\left(\frac{\Delta E\, t}{2\hbar}\right)$$

Ce phénomène de transformation est appelé *oscillation neutrino.*

3. Si $p \gg \langle m \rangle c$ est l'impulsion des neutrinos, montrer que

$$\Delta E = \frac{(m_2^2 - m_1^2)c^3}{2p} = \frac{\Delta m^2\, c^3}{2p}$$

avec $\Delta m^2 = m_2^2 - m_1^2$.

4. En supposant qu'il existe une demi-oscillation sur le parcours Soleil-Terre (c'est-à-dire $\Delta E\, t/\hbar = \pi$) pour des neutrinos de 8 MeV, quel est l'ordre de propriété de la différence des masses carrées Δm^2 ? La distance Terre-Soleil est de 150 millions de kilomètres.

4.4.7 Points de vue de Schrödinger et de Heisenberg

Soit un opérateur hermitien A dépendant du temps dans le point de vue de Schrödinger : $A = A(t)$. Le hamiltonien H est aussi supposé dépendant du temps. Montrer que

$$A_{\rm H}(t) = U^{-1}(t, t_0)A(t)U(t, t_0)$$

vérifie

$$\mathrm{i}\hbar \frac{\mathrm{d}A_{\rm H}}{\mathrm{d}t} = [A_{\rm H}(t), H_{\rm H}(t)] + \mathrm{i}\hbar \left(\frac{\partial A(t)}{\partial t}\right)_{\rm H}$$

où $H_{\rm H}(t)$ et $(\partial A/\partial t)_{\rm H}$ sont obtenus à partir de $H(t)$ et $(\partial A(t)/\partial t)$ par la loi de transformation utilisée pour A.

4.4.8 Borne de Helstrom

On prépare un grand nombre d'états quantiques $|\varphi\rangle$ et $|\chi\rangle$, l'état $|\varphi\rangle$ avec la probabilité r, l'état $|\chi\rangle$ avec la probabilité $s = 1 - r$. Un expérimentateur effectue une mesure sur ces états avec l'objectif de distinguer les états $|\varphi\rangle$ et $|\chi\rangle$. Le but de l'exercice est de déterminer la stratégie optimale pour effectuer cette distinction et le taux de succès de cette stratégie. En plus de son intérêt théorique, le résultat obtenu, la borne de Helstrom, a des application pratiques en informatique quantique. Une autre façon d'aborder le problème est l'objet de l'exercice 11.6.13.

Si les états $|\varphi\rangle$ et $|\chi\rangle$ sont orthogonaux, il suffit de mesurer le projecteur $|\varphi\rangle\langle\varphi|$: si le résultat de la mesure est 1 (0), alors c'est l'état $|\varphi\rangle$ ($|\chi\rangle$) qui a été préparé. Dans le cas général, on cherche un projecteur $\mathcal{P}$ en décidant *a priori* que $\mathcal{P} = 1$ ($\mathcal{P} = 0$) correspond au résultat $|\varphi\rangle$ ($|\chi\rangle$). Ce projecteur doit donc maximiser la quantité

$$X = r\langle\varphi|\mathcal{P}|\varphi\rangle + s\langle\chi|\mathcal{P}|\chi\rangle$$

où, sans perte de généralité, $r \geq s$. Dans un premier temps, on se place dans l'espace à deux dimensions sous-tendu par $|\varphi\rangle$ et $|\chi\rangle$. On montrera à la question **2** que cela ne restreint pas la généralité. Posant $\mathcal{P} = |\psi\rangle\langle\psi|$, on écrit $|\varphi\rangle$, $|\chi\rangle$ et $|\psi\rangle$ sous la forme

$$|\varphi\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |\chi\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}\,\mathrm{e}^{\,\mathrm{i}\phi} \end{pmatrix} \qquad |\psi\rangle = \begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2}\,\mathrm{e}^{\,\mathrm{i}\beta} \end{pmatrix}$$

et on choisit dans un premier temps $\phi = \beta = 0$.

1. Montrer que

$$2X - 1 = r\cos\alpha - s\cos(\alpha - \theta)$$

En cherchant le maximum de X par rapport à α, montrer que

$$\sin\alpha = -\frac{\sin\theta}{\sqrt{D}} \qquad \cos\alpha = \frac{r - s\cos\theta}{\sqrt{D}} \qquad D = 1 - 4rs\cos^2\frac{\theta}{2}$$

En déduire le taux de succès maximum lorsque l'on cherche à faire la distinction entre $|\varphi\rangle$ et $|\chi\rangle$

$$\mathsf{p}_{\max} = \frac{1}{2}\left(1 + \sqrt{D}\right) = \frac{1}{2}\left(1 + \sqrt{1 - 4rs|\langle\varphi|\chi\rangle|^2}\right)$$

ce qui est la *borne de Helstrom*. En déduire que $S = |\langle\chi|\varphi\rangle|$ peut être considéré comme une mesure quantitative de la discernabilité de $|\varphi\rangle$ et $|\chi\rangle$.

2. Montrer que le résultat n'est pas affecté si l'on tient compte des angles ϕ et β, et qu'il reste aussi valable si l'on ne se restreint pas pour $|\psi\rangle$ à l'espace à deux dimensions sous-tendu par $|\varphi\rangle$ et $|\chi\rangle$.

4.4.9 Règle de Born généralisée

On considère un état à deux particules A et B (voir la section 11.1 pour plus de détails sur ce type d'états appelés états intriqués)

$$|\Psi\rangle = \sum_{i,j} c_{ij}|\varphi_i \otimes \chi_j\rangle \qquad i = 1, \cdots, N_A \qquad j = 1, \cdots, N_B$$

où $|\varphi_i\rangle$ ($|\chi_j\rangle$) désigne un état de la particule A (B). On mesure le projecteur

$$\mathcal{P}_{i_0 j_0} = |\varphi_{i_0} \otimes \chi_{j_0}\rangle\langle\varphi_{i_0} \otimes \chi_{j_0}|$$

Quelle est la probabilité de trouver $\mathcal{P}_{i_0 j_0} = 1$? Supposons que l'on effectue d'abord une mesure sur la particule B et qu'on la trouve dans l'état $|\chi_{j_0}\rangle$. Montrer que le vecteur d'état de la particule A après cette mesure est

$$|\varphi\rangle = \frac{1}{\sqrt{\mathsf{p}(j_0)}}\sum_i c_{ij_0}|\varphi_i\rangle \qquad \mathsf{p}(j_0) = \sum_i |c_{ij_0}|^2$$

Quelle est la probabilité conditionnelle $\mathsf{p}(i_0|j_0)$ de trouver la particule A dans l'état $|\varphi_{i_0}\rangle$ sachant que la particule B a été trouvée dans l'état $|\chi_{j_0}\rangle$? Montrer que cette probabilité obéit bien à la loi de Bayes pour les probabilités conditionnelles.

4.4.10 Le système des mésons K neutres : évolution non unitaire

Supposons que l'on crée au temps $t = 0$ une particule instable A de masse m. Le vecteur d'état au temps $t = 0$ est $|\varphi\rangle$. Au temps t, la composante $c(t)$ de $|\psi(t)\rangle$ sur $|\varphi\rangle$ est $c(t) = \langle\varphi|\psi(t)\rangle$ (voir (4.31)). Si la particule A était stable, $c(t)$ serait simplement donné par

$$c(t) = \exp\Big(-\mathrm{i}\frac{Et}{\hbar}\Big) = \exp\Big(-\mathrm{i}\frac{mc^2t}{\hbar}\Big)$$

dans son référentiel au repos, où son énergie est $E = mc^2$, et l'on aurait $|c(t)|^2 = 1$ pour tout t : la probabilité que la particule existe au temps t serait toujours égale à un. Cependant, on suppose la particule instable et sa loi de désintégration suit une exponentielle (4.36)

$$c(t) = \exp\Big(-\mathrm{i}\frac{mc^2t}{\hbar}\Big)\exp\Big(-\frac{t}{2\tau}\Big)$$

où τ est la vie moyenne. Si l'on s'intéresse *uniquement* au vecteur d'état $|\varphi(t)\rangle$ de la particule, on peut rendre compte de la forme de $c(t)$ en utilisant une *évolution non unitaire* (ceci sera vu en détail au chapitre 18 : l'évolution d'un système qui n'est pas fermé n'est pas unitaire) et écrire une équation diférentielle pour $c(t)$

$$\mathrm{i}\hbar\dot{c}(t) = \Big(mc^2 - \mathrm{i}\,\frac{\hbar\Gamma}{2}\Big)c(t)$$

On se propose de généraliser cette description au cas d'un système à deux niveaux, le système des mésons K neutres, écrite dans le référentiel au repos de ces mésons : t est le temps propre, en généralisant l'équation précédente. Il existe deux types de mésons K neutres[22], le méson K^0 formé d'un quark d et d'un antiquark étrange $\overline{s}$, et le méson $\overline{K^0}$ formé d'un antiquark $\overline{d}$ et d'un quark étrange s. On rappelle que les charges des quarks u, d et s sont respectivement $2/3$, $-1/3$ et $-1/3$ en unité de la charge du proton. La production de ces mésons se fait par interaction forte, et cette interaction vérifie une loi de conservation analogue à celle de la charge électrique : le nombre de quarks étranges moins le nombre d'antiquarks étranges est conservé (de même que dans une réaction impliquant des électrons et des positrons, le nombre d'électrons moins le nombre de positrons est conservé par conservation de la charge

22. Il existe aussi deux mésons K chargés, le méson K^+ ($u\overline{s}$) et le méson K^- ($\overline{u}s$).

électrique). Donnons quelques exemples : le méson π^+ est une combinaison $(u\,\overline{d})$, le méson π^- une combinaison $(\overline{u}\,d)$ et la particule Λ^0 une combinaison (uds). Les réactions

$$\pi^-(\overline{u}\,d) + \text{proton }(uud) \to K^0(d\,\overline{s}) + \Lambda^0\ (uds)$$

ou

$$\overline{K^0}\,(\overline{d}\,s) + \text{proton}\,(uud) \to \pi^+\,(u\,\overline{d}) + \Lambda^0\,(uds)$$

sont permises, tandis que

$$\pi^-\,(\overline{u}d) + \text{proton}\,(uud) \to \overline{K^0}(\overline{d}\,s) + \Lambda^0\ (uds)$$

ou

$$K^0\,(d\,\overline{s}) + \text{proton}\,(uud) \to \pi^+(u\,\overline{d}) + \Lambda^0\,(uds)$$

sont interdites.

1. Le système $(K^0, \overline{K^0})$ est un système à deux niveaux dont le vecteur d'état $|\varphi(t)\rangle$ peut s'écrire

$$|\varphi(t)\rangle = c(t)|K^0\rangle + \overline{c}(t)|\overline{K^0}\rangle$$

dans la base $\{|K^0\rangle, |\overline{K^0}\rangle\}$. Les composantes du vecteur $|\varphi(t)\rangle$ obéissent à une équation d'évolution

$$\mathrm{i}\hbar\begin{pmatrix}\dot{c}(t)\\ \dot{\overline{c}}(t)\end{pmatrix} = M\begin{pmatrix}c(t)\\ \overline{c}(t)\end{pmatrix}$$

où M est un matrice 2×2. Soit $\mathcal{C}$ l'opérateur de "conjugaison de charge" qui échange particules et antiparticules[23]

$$\mathcal{C}|K^0\rangle = |\overline{K^0}\rangle \qquad \mathcal{C}|\overline{K^0}\rangle = |K^0\rangle$$

Montrer que si M commute avec $\mathcal{C}$, sa forme la plus générale est

$$M = \begin{pmatrix} A & B\\ B & A\end{pmatrix}$$

où A et B sont *a priori* deux nombres complexes, car la matrice M n'est pas hermitienne.

2. Quels sont les vecteurs propres $|K_1\rangle$ et $|K_2\rangle$ de M ? Montrer que ce sont ces deux états qui ont une énergie et une vie moyenne bien déterminées. Si, au temps $t=0$, $|\varphi(t)\rangle$ a pour composantes $c(0)$ et $\overline{c}(0)$, calculer $c(t)$ et $\overline{c}(t)$. On posera

$$\begin{aligned} A &= \frac{1}{2}\left[(E_1+E_2) - \frac{\mathrm{i}\hbar}{2}\,(\Gamma_1+\Gamma_2)\right]\\ B &= \frac{1}{2}\left[(E_1-E_2) - \frac{\mathrm{i}\hbar}{2}\,(\Gamma_1-\Gamma_2)\right]\end{aligned}$$

23. On peut généraliser le raisonnement en utilisant au lieu de $\mathcal{C}$ le produit $\mathcal{CP}$, où $\mathcal{P}$ est l'opération parité. En fait, l'expérience montre que $[M, \mathcal{CP}] \neq 0$, mais les corrections sont très faibles.

3. On produit au temps $t = 0$ un méson K^0 dans la réaction

$$\pi^-(\overline{u}\,d) + \text{proton}\ (uud) \to K^0(d\overline{s}) + \Lambda^0\ (uds)$$

Quelle est la probabilité[24] pour trouver un méson $\overline{K^0}$ au temps t ? En supposant $\Gamma_1 \gg \Gamma_2$, montrer que la probabilité d'observer la réaction

$$\overline{K^0}\,(\overline{d}\,s) + \text{proton}\,(uud) \to \pi^+\,(u\,\overline{d}) + \Lambda^0\,(uds)$$

est proportionnelle pour $t \sim \tau_1 = 1/\Gamma_1$ à

$$\mathsf{p}(t) = 1 - 2\exp\left(-\frac{\Gamma_1\,t}{2}\right)\cos\frac{(E_1 - E_2)t}{\hbar} + \exp\left(-\Gamma_1\,t\right)$$

Tracer la courbe représentative de $\mathsf{p}(t)$. Que pensez-vous des ordres de grandeur respectifs de $(E_1 - E_2)$ et de E_1 ou E_2 ? Comment pourrait-on mesurer $(E_1 - E_2)$? Les valeurs numériques sont : $\tau_1 \simeq 10^{-10}\,\text{s}$, $\tau_2 \simeq 10^{-7}\,\text{s}$, $E_1 \simeq E_2 \simeq 500\,\text{MeV}$.

4.5 Bibliographie

Notre présentation des postulats de la mécanique quantique s'écarte sensiblement de celle des exposés classiques que l'on trouvera par exemple dans Messiah [1959], chapitre VIII, Cohen-Tannoudji *et al.* [1973], chapitre III, ou Basdevant et Dalibard [2001], chapitre 5. Le lecteur pourra aussi consulter Ballentine [1998], chapitre 9, Peres [1993], chapitre 2, Isham [1995], chapitre 5 et Omnès [1994]. Une discussion qualitative des inégalités de Heisenberg se trouve dans Lévy-Leblond et Balibar [1984], chapitre 3. Les inégalités de Heisenberg temporelles sont traitées de façon imprécise, voire incorrecte, dans certains manuels ; pour une discussion approfondie, voir par exemple Peres [1993], chapitre 12, et Ballentine [1998], chapitre 12.

24. En pratique, les mésons K se propagent en ligne droite à partir de leur point de production avec une vitesse proche de la vitesse de la lumière, et on se place à une distance $l \simeq ct(1 - v^2/c^2)^{-1/2}$ du point de production.

Chapitre 5

Systèmes à nombre de niveaux fini

Ce chapitre examine quelques applications simples de la mécanique quantique, dans des situations où l'on obtient une bonne modélisation d'un système quantique en se restreignant à un espace des états de dimension finie. Si chaque niveau d'énergie, même dégénéré, est compté une fois, la dimension de l'espace des états $\mathcal{H}$ est égale au nombre de niveaux : c'est pourquoi on parle de *système à nombre de niveaux fini*. Les deux premiers exemples seront empruntés à la chimie quantique et nous permettront d'étudier le cas stationnaire, celui d'un hamiltonien indépendant du temps. Cependant, le point essentiel de ce chapitre est l'introduction de la dépendance par rapport au temps induite par un couplage d'un système à deux niveaux avec un champ extérieur classique périodique. Cette dépendance par rapport au temps sera illustrée par trois exemples d'une grande importance pratique : la résonance magnétique nucléaire ou RMN (section 5.2), la molécule d'ammoniac (section 5.3) et l'atome à deux niveaux (section 5.4).

5.1 Chimie quantique élémentaire

5.1.1 Molécule d'éthylène

La molécule d'éthylène C_2H_4 servira d'introduction au sujet. L'"ossature" de cette molécule est formée par les liaisons dites *liaisons* σ : des paires *d'électrons* σ de spin opposé sont mises en commun entre les deux atomes de carbone ainsi qu'entre les atomes de carbone et les atomes d'hydrogène, formant un ion $(C_2H_4)^{++}$ (figure 5.1). Il reste à placer deux électrons, appelés *électrons* π, qui sont mobiles : schématiquement ces deux électrons peuvent sauter d'un atome de carbone à l'autre. On dit qu'ils sont *délocalisés*. Le fait de traiter séparément électrons π et électrons σ est bien sûr une approximation, mais cette approximation joue un grand rôle dans la théorie de la liaison chimique. Commençons par placer le premier électron π. Celui-ci peut être localisé au voisinage de l'atome de carbone 1 ; l'état quantique correspondant sera

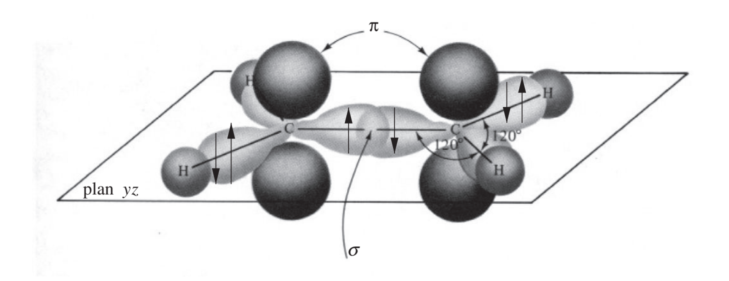


FIG. 5.1 – La molécule d'éthylène.

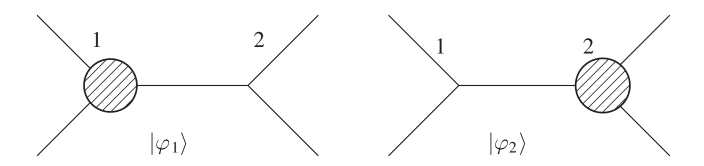


FIG. 5.2 – Les deux états possibles d'un électron π, localisés au voisinage de l'atome 1 ou de l'atome 2.

désigné par $|\varphi_1\rangle$. Il peut aussi être localisé au voisinage de l'atome de carbone 2, et l'état quantique correspondant sera désigné par $|\varphi_2\rangle$ (figure 5.2). L'énergie de cet électron localisé sur l'atome 1 ou l'atome 2 est E_0, la même par symétrie entre les deux atomes. Nous allons prendre comme approximation de l'espace des états un espace à deux dimensions $\mathcal{H}$ dont les vecteurs de base sont $\{|\varphi_1\rangle, |\varphi_2\rangle\}$. Dans cette base, le hamiltonien s'écrit provisoirement

$$H_0 = \begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix} \qquad H_0|\varphi_{1,2}\rangle = E_0|\varphi_{1,2}\rangle \tag{5.1}$$

Cependant, ce hamiltonien est incomplet, car nous n'avons pas tenu compte de la possibilité pour l'électron de sauter d'un atome de carbone à l'autre. Dans le cadre de nos approximations, qui sont celles de la théorie des *orbitales moléculaires de Hückel*, la forme la plus générale de H est

$$H = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \tag{5.2}$$

et l'élément non diagonal $-A$ de H autorise précisément des transitions entre $|\varphi_1\rangle$ et $|\varphi_2\rangle$. Un choix adéquat de la phase des vecteurs de base nous a permis de prendre A réel : *cf.* le § 2.3.2. Nous avons affecté A d'un signe $(-)$ qui n'est pas indifférent, car il est possible de montrer que $A > 0$.

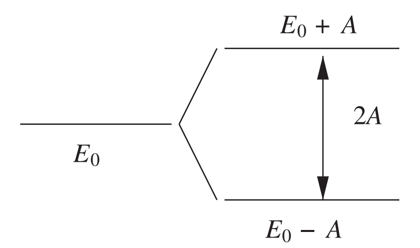


FIG. 5.3 – Niveaux d'énergie d'un électron π.

Si $A \neq 0$, les états $|\varphi_1\rangle$ et $|\varphi_2\rangle$ ne sont plus des états stationnaires. Comme nous l'avons vu au § 2.3.2, les vecteurs propres de H sont maintenant

$$|\chi_+\rangle = \frac{1}{\sqrt{2}}\Big(|\varphi_1\rangle + |\varphi_2\rangle\Big) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \tag{5.3}$$

$$|\chi_-\rangle = \frac{1}{\sqrt{2}}\Big(|\varphi_1\rangle - |\varphi_2\rangle\Big) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} \tag{5.4}$$

avec

$$H|\chi_+\rangle = (E_0 - A)|\chi_+\rangle \qquad H|\chi_-\rangle = (E_0 + A)|\chi_-\rangle \tag{5.5}$$

Comme $A > 0$, l'état symétrique $|\chi_+\rangle$ est l'état d'énergie la plus basse. Le spectre du hamiltonien est donné sur la figure 5.3 : l'état fondamental est l'état $|\chi_+\rangle$, d'énergie $(E_0 - A)$. On peut donner une interprétation spatiale de ces résultats en examinant la localisation de l'électron sur la droite joignant les deux atomes de carbone prise comme axe des x, l'origine étant située au milieu de cette droite. Comme nous le verrons en détail au chapitre 8, si $|x\rangle$ est un vecteur propre de l'opérateur position, la quantité $\langle x|\varphi_1\rangle$ est l'amplitude de probabilité pour trouver au point x l'électron dans l'état $|\varphi_1\rangle$. Au chapitre 8, cette amplitude de probabilité sera appelée la *fonction d'onde* de l'électron. Le module au carré de cette amplitude de probabilité donne la probabilité[1] de trouver l'électron au point x, aussi appelée *probabilité de présence de l'électron au point x*. Cette interprétation permet de représenter qualitativement sur la figure 5.4 les amplitudes de probabilité $\chi_\pm(x) = \langle x|\chi_\pm\rangle$ correspondant aux états $|\chi_\pm\rangle$. La probabilité de présence correspondante s'annule à l'origine dans le cas antisymétrique $|\chi_-\rangle$, mais non dans le cas symétrique $|\chi_+\rangle$. Le caractère symétrique ou antisymétrique de la fonction d'onde de l'état fondamental est lié au signe de A. En pratique, un état fondamental est toujours symétrique, ce qui correspond à $A > 0$.

Il nous reste à placer le second électron : ceci se fera très simplement si nous pouvons ignorer les interactions entre cet électron et le précédent, c'est-à-dire utiliser l'approximation des électrons indépendants. Pour obtenir l'état

1. Plus précisément, c'est une probabilité par unité de longueur : $|\langle x|\varphi\rangle|^2 \mathrm{d}x$ est la probabilité de trouver la particule dans l'intervalle $[x, x + \mathrm{d}x]$: voir le § 8.1.2.

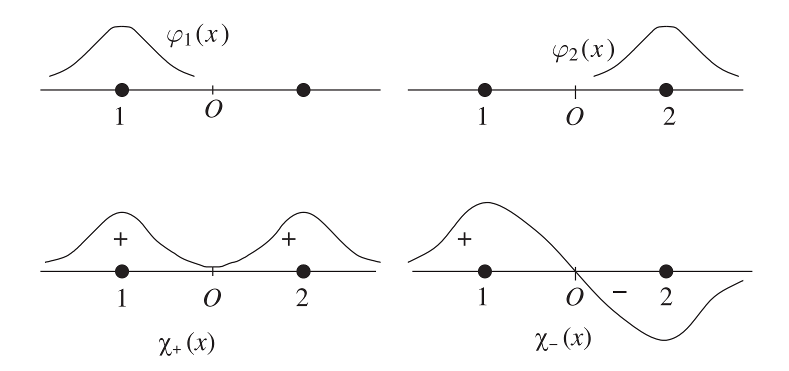


FIG. 5.4 – Amplitudes de probabilité pour trouver un électron π en un point x.

fondamental, il suffit de placer le second électron dans l'état $|\chi_+\rangle$ d'énergie $(E_0 - A)$. Le principe de Pauli (chapitre 14) restreint alors les états de spin : si le premier électron a son spin up ($|+\rangle$), le second électron doit avoir son spin down ($|-\rangle$), ainsi que nous le verrons au chapitre 14. L'état fondamental de la liaison π est finalement $2(E_0 - A)$; $-2A$ est appelé *énergie de délocalisation* des électrons π. Il faut souligner le rôle crucial de *l'approximation des particules indépendantes* utilisée dans le raisonnement ci-dessus : les électrons π n'interagissent pas avec les électrons σ et n'interagissent pas entre eux. Cette modélisation est difficile à justifier à partir des principes fondamentaux, ou de ce que l'on appelle aujourd'hui des calculs *ab initio*, mais elle se révèle d'un intérêt pratique considérable.

5.1.2 Molécule de benzène

Dans la molécule de benzène, l'ossature σ de l'ion $(C_6H_6)^{6+}$ forme un hexagone. Si l'on rajoute les 6 électrons π de façon à former trois doubles liaisons, on obtient la formule de Kékulé (figure 5.5a), et on prédit une énergie $6(E_0 - A)$ pour l'état fondamental. On sait par un raisonnement de chimie que la formule de Kékulé ne peut pas être tout à fait correcte[2], et nous allons effectivement voir que tenir compte de la délocalisation des électrons π tout au long de la chaîne hexagonale conduit à une énergie plus basse que $6(E_0 - A)$: la formule de Kékulé ne donne pas correctement l'énergie de l'état fondamental. Examinons pour commencer l'addition d'un seul électron, et numérotons[3] de 0 à 5 les atomes de carbone le long de la chaîne hexagonale en prenant une origine arbitraire (figure 5.5b). Nous notons $|\varphi_3\rangle$ par exemple l'état où l'élec-

2. Par exemple, il existe une seule forme d'orthodibromobenzène, alors que la formule de Kékulé en prévoit deux différentes. On peut aussi remarquer que la longueur de la liaison entre deux atomes de carbone dans le benzène (1.40 Å) est intermédiaire entre une simple lisaison (1.54 Å) et une double liaison (1.35 Å).

3. Comme nous le verrons dans un instant, il est plus commode de numéroter de 0 à 5 que de 1 à 6 !

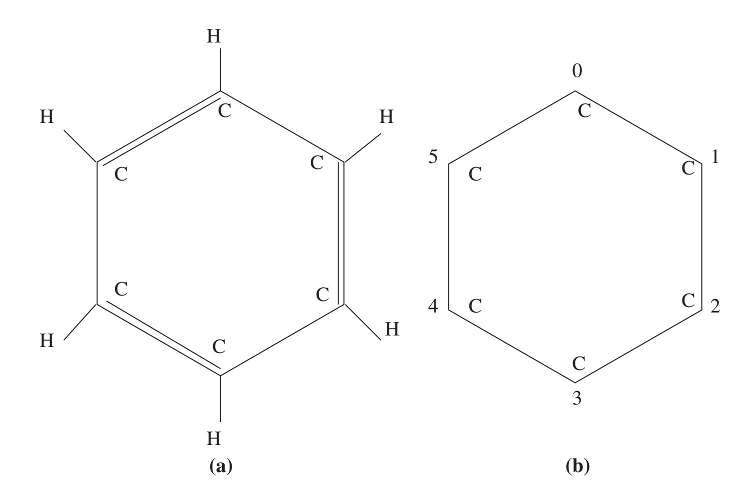


FIG. 5.5 – (a) Configuration hexagonale d'une molécule de benzène et formule de Kékulé. (b) Ossature des électrons σ.

tron est localisé au voisinage de l'atome n° 3. Comme il n'est pas plus difficile de traiter un nombre quelconque N d'atomes de carbone formant une chaîne fermée, c'est-à-dire un polygone régulier à N côtés, nous notons $|\varphi_n\rangle$ l'état où l'électron est localisé au voisinage de l'atome n° n, $n = 0,\ 1,\ \ldots,\ N-1$, en prenant $N = 6$ pour le benzène. Les atomes n et $n + N$ sont identiques : $n \equiv n+N$. L'espace des états est à N dimensions, et le hamiltonien est défini par son action sur $|\varphi_n\rangle$

$$H|\varphi_n\rangle = E_0|\varphi_n\rangle - A(|\varphi_{n-1}\rangle + |\varphi_{n+1}\rangle) \tag{5.6}$$

Pour trouver les valeurs propres et vecteurs propres de H, nous allons exploiter la symétrie du problème sous toute permutation circulaire des N atomes de la chaîne. Soit U_P l'opérateur unitaire qui effectue une permutation circulaire des atomes dans le sens $n \to (n-1)$

$$U_P|\varphi_n\rangle = |\varphi_{n-1}\rangle \qquad U_P^\dagger|\varphi_n\rangle = U_P^{-1}|\varphi_n\rangle = |\varphi_{n+1}\rangle \tag{5.7}$$

D'après (5.6) et (5.7), nous pouvons écrire le hamiltonien sous la forme

$$H = E_0 I - A(U_P + U_P^\dagger) \tag{5.8}$$

ce qui implique que H et U_P commutent

$$[H, U_P] = 0 \tag{5.9}$$

et ont une base de vecteurs propres communs. Cherchons les vecteurs propres et valeurs propres de U_P, qui est *a priori* un opérateur plus simple que H.

Comme U_P est unitaire, ses valeurs propres sont de la forme $\exp(\mathrm{i}\delta)$ (*c.f.* le § 2.3.4). Comme $(U_P)^N = I$, on doit avoir $\exp(\mathrm{i}N\delta) = 1$, et par conséquent les valeurs propres sont indicées par un indice entier s

$$\delta = \delta_s = \frac{2\pi s}{N} \qquad s = 0, 1, \ \ldots, \ N-1 \tag{5.10}$$

Nous avons donc déterminé N valeurs propres distinctes de U_P. Comme U_P agit dans un espace de dimension N, les vecteurs propres correspondants sont orthogonaux et forment une base de $\mathcal{H}$. Écrivons un vecteur propre normalisé $|\chi_s\rangle$ sous la forme

$$|\chi_s\rangle = \sum_{n=0}^{N-1} c_n |\varphi_n\rangle \qquad \sum_{n=0}^{N-1} |c_n|^2 = 1 \tag{5.11}$$

Nous avons, d'une part,

$$U_P|\chi_s\rangle = \sum_{n=0}^{N-1} c_n |\varphi_{n-1}\rangle = \sum_{n=0}^{N-1} c_{n+1} |\varphi_n\rangle$$

et, d'autre part,

$$U_P|\chi_s\rangle = \mathrm{e}^{\mathrm{i}\delta_s}|\chi_s\rangle = \sum_{n=0}^{N-1} \mathrm{e}^{\mathrm{i}\delta_s} c_n |\varphi_n\rangle$$

L'identification des coefficients de $|\varphi_n\rangle$ dans ces deux équations conduit à

$$c_{n+1} = \mathrm{e}^{\,\mathrm{i}\delta_s} c_n \quad \text{soit} \quad c_n = \mathrm{e}^{\,\mathrm{i}n\delta_s} c_0$$

Ceci donne pour le vecteur propre correspondant à la valeur propre $\exp(\mathrm{i}\delta_s)$

$$|\chi_s\rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathrm{e}^{\mathrm{i}n\delta_s}\, |\varphi_n\rangle \tag{5.12}$$

Le choix $c_0 = 1/\sqrt{N}$ assure la normalisation de $|\chi_s\rangle$. Compte tenu de l'expression (5.8) de H, la valeur propre E_s est donnée par

$$E_s = E_0 - A\Big(\mathrm{e}^{\,\mathrm{i}\delta_s} + \mathrm{e}^{-\mathrm{i}\delta_s}\Big) = E_0 - 2A\cos\delta_s$$

soit (figure 5.6)

$$E_s = E_0 - 2A\cos\frac{2\pi s}{N} \tag{5.13}$$

Nous aurions pu arriver directement à (5.13) sans passer par l'intermédiaire de l'opérateur de permutation circulaire U_P. Néanmoins, ce passage par U_P illustre une stratégie générale et non une "astuce de calcul". Nous aurons souvent à mettre en œuvre cette stratégie, car elle simplifie, et parfois de

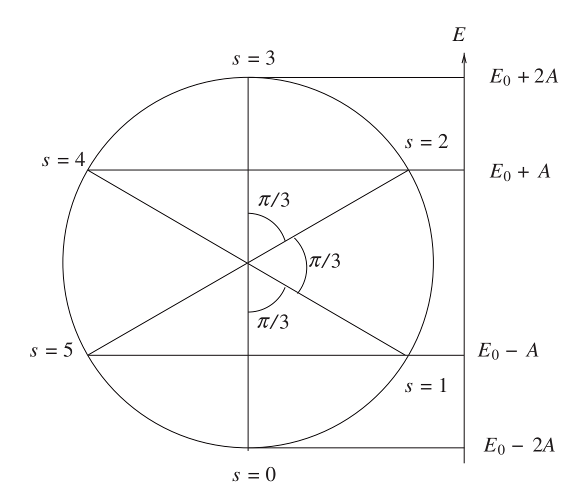


FIG. 5.6 – Niveaux d'énergie d'un électron π de la molécule de benzène.

façon considérable, la diagonalisation du hamiltonien : au lieu de diagonaliser directement H, on diagonalise d'abord les opérateurs de symétrie unitaires qui commutent avec H, lorsque de tels opérateurs existent en raison d'une symétrie du problème physique.

On remarque que les valeurs s et $\tilde{s} = N - s$ donnent les mêmes valeurs de l'énergie : en dehors de $s = 0$ et $s = N - 1$ (pour N pair), les niveaux d'énergie sont deux fois dégénérés. Il est possible d'écrire les vecteurs propres de H avec des composantes réelles en prenant des combinaisons linéaires de $|\chi_s\rangle$ et $|\chi_{\tilde{s}}\rangle$

$$|\chi_s^+\rangle = \frac{1}{\sqrt{2}}\left(|\chi_s\rangle + |\chi_{\tilde{s}}\rangle\right) = \sqrt{\frac{2}{N}}\sum_{n=0}^{N-1}\cos\frac{2\pi ns}{N}\,|\varphi_n\rangle \tag{5.14}$$

$$|\chi_s^-\rangle = \frac{1}{\mathrm{i}\sqrt{2}}\left(|\chi_s\rangle - |\chi_{\tilde{s}}\rangle\right) = \sqrt{\frac{2}{N}}\sum_{n=0}^{N-1}\sin\frac{2\pi ns}{N}\,|\varphi_n\rangle \tag{5.15}$$

Nous pouvons maintenant rassembler les résultats pour les valeurs propres de H et les vecteurs propres correspondants dans le cas du benzène : $N = 6$, $\cos(2\pi/6) = 1/2$, $\sin(2\pi/6) = \sqrt{3}/2$ (figure 5.6)

$$\begin{aligned}
s &= 0 \qquad\qquad E = E_0 - 2A \\
|\chi_0\rangle &= \frac{1}{\sqrt{6}}(1,1,1,1,1,1) \\
s &= 1,\ \tilde{s} = 5 \qquad\qquad E = E_0 - A \\
|\chi_1^+\rangle &= \frac{1}{\sqrt{3}}\left(1,\frac{1}{2},-\frac{1}{2},-1,-\frac{1}{2},\frac{1}{2}\right) \qquad\qquad |\chi_1^-\rangle = \left(0,\frac{1}{2},\frac{1}{2},0,-\frac{1}{2},-\frac{1}{2}\right)
\end{aligned}$$

$$
\begin{aligned}
s &= 2,\ \tilde{s} = 4 \qquad E = E_0 + A \\
|\chi_2^+\rangle &= \frac{1}{\sqrt{3}}\left(1, -\frac{1}{2}, -\frac{1}{2}, 1, -\frac{1}{2}, -\frac{1}{2}\right) \qquad |\chi_2^-\rangle = \left(0, \frac{1}{2}, -\frac{1}{2}, 0, \frac{1}{2}, -\frac{1}{2}\right) \\
s &= \tilde{s} = 3 \qquad E = E_0 + 2A \\
|\chi_3\rangle &= \frac{1}{\sqrt{6}}(1, -1, 1, -1, 1, -1)
\end{aligned}
\tag{5.16}
$$

Cherchons maintenant l'état fondamental, c'est-à-dire l'état de plus basse énergie, en plaçant les 6 électrons π délocalisés. À l'approximation des électrons indépendants, cet état sera obtenu en mettant d'abord deux électrons de spin opposé dans le niveau $E_0 - 2A$, le principe de Pauli (chapitre 14) nous interdisant d'y mettre d'autres électrons. Le niveau $(E_0 - A)$ étant doublement dégénéré, nous pouvons y mettre quatre électrons (deux paires d'électrons de spin opposé) ce qui donne une énergie totale

$$
E = 2(E_0 - 2A) + 4(E_0 - A) = 6E_0 - 8A \tag{5.17}
$$

Cette énergie est inférieure de $2A$ à celle $(6E_0 - 6A)$ de la formule de Kékulé : les électrons π du benzène ne sont pas localisés sur des doubles liaisons, mais ils sont délocalisés le long de l'ensemble de la chaîne hexagonale, et cette forme de délocalisation diminue l'énergie de $2A$.

La comparaison entre la chaleur[4] d'hydrogénation du benzène en cyclohexane

$$
C_6H_6 + 3H_2 \to C_6H_{12} - 49.8 \text{ kcal/mole}
$$

et celle du cyclohexène, qui contient une seule double liaison

$$
C_6H_{10} + H_2 \to C_6H_{12} - 28.6 \text{ kcal/mole}
$$

permet d'estimer $2A$: $2A = 3 \times 28.6 - 49.8 = 36$ kcal/mole $\simeq 1.6\,\text{eV}$. Cependant, cette estimation est au mieux un ordre de grandeur, car elle est sujette à des incertitudes difficiles à apprécier. Elles sont dues à l'approximation des électrons indépendants, qui est loin d'être bien contrôlée.

5.2 Résonance magnétique nucléaire (RMN)

Dans la section 5.1, nous avons étudié les niveaux d'énergie de hamiltoniens indépendants du temps. Dans les trois sections qui vont suivre, nous allons introduire une interaction dépendant du temps pour un système à deux niveaux, en plongeant le système dans un champ extérieur classique périodique de fréquence ω. Dans ces conditions, il n'y a évidemment plus d'états stationnaires, et *le problème intéressant devient l'étude des transitions d'un niveau à l'autre sous l'influence du champ extérieur*. Nous montrerons le résultat fondamental

4. Pour les puristes : il s'agit en fait d'une variation d'enthalpie, mais la différence est négligeable.

suivant : si $\omega \simeq \omega_0$, où $\hbar\omega_0$ est la différence d'énergie entre les deux niveaux, on observe un remarquable *phénomène de résonance*. Nous allons en donner trois exemples d'une grande importance pratique : la résonance magnétique nucléaire dans cette section, la molécule d'ammoniac dans la section 5.3 et l'atome à deux niveaux dans la section 5.4.

5.2.1 Spin 1/2 dans un champ magnétique périodique

La résonance magnétique nucléaire (RMN) repose sur le fait que les noyaux atomiques de spin non nul possèdent des moments magnétiques. Nous nous limiterons aux noyaux de spin 1/2 (^{1}H, ^{13}C, ^{19}F, etc.) dont le moment magnétique, qui est un opérateur en mécanique quantique, est donné par

$$\vec{\mu} = \gamma \vec{S} = \frac{1}{2}\,\gamma\hbar\vec{\sigma} \tag{5.18}$$

où $\vec{S}$ est l'opérateur de spin défini dans la section 3.2, γ est le facteur gyromagnétique

$$\gamma = \overline{\gamma}\,\frac{q_p}{2m_p} \tag{5.19}$$

avec $\overline{\gamma} = 5.59$ pour le proton, 1.40 pour le ^{13}C, 5.26 pour le ^{19}F, etc. Le spin nucléaire est placé dans un champ magnétique $\vec{B}_0$ dirigé suivant Oz. Suivant (3.61), le hamiltonien H_0 du spin nucléaire s'écrit

$$H_0 = -\vec{\mu}\cdot\vec{B}_0 = -\frac{1}{2}\,\gamma\hbar B_0\sigma_z = -\frac{1}{2}\,\hbar\omega_0\sigma_z \tag{5.20}$$

avec $\omega_0 = \gamma B_0$, soit encore sous forme matricielle dans une base où σ_z est diagonal

$$H_0 = -\frac{1}{2}\,\hbar\begin{pmatrix} \omega_0 & 0 \\ 0 & -\omega_0 \end{pmatrix} \tag{5.21}$$

On note que comme la charge du proton q_p est positive, on n'introduit pas de signe moins dans la définition de ω_0, contrairement à ce qui avait été fait dans la section 3.2.5 pour le cas de l'électron. ω_0 est la fréquence de Larmor, la fréquence de précession du moment magnétique classique autour de $\vec{B}_0$ (figure 3.7) ; dans le cas du proton la précession de Larmor s'effectue dans le sens inverse du sens trigonométrique. L'état $|+\rangle$ a une énergie $-\hbar\omega_0/2$, l'état $|-\rangle$ une énergie $\hbar\omega_0/2$; on est donc en présence d'un système à deux niveaux : les deux niveaux Zeeman (§ 3.2.5) d'un spin 1/2 dans un champ magnétique, la différence d'énergie étant $\hbar\omega_0$.

On ajoute au champ constant $\vec{B}_0$ un champ périodique $\vec{B}_1(t)$ situé dans le plan xOy tournant dans le sens inverse du sens trigonométrique[5], c'est-à dire dans le même sens que la précession de Larmor, avec une vitesse angulaire ω

$$\vec{B}_1(t) = B_1(\hat{x}\cos\omega t - \hat{y}\sin\omega t) \tag{5.22}$$

5. On pourrait aussi prendre un champ $\vec{B}_1(t)$ parallèle à Ox : *cf.* l'exercice 5.5.6.

En pratique, un tel champ peut être obtenu au moyen de deux bobines placées le long des axes Ox et Oy, alimentées en courant alternatif de fréquence ω. La contribution au hamiltonien induite par le champ $\vec{B}_1(t)$ est

$$H_1(t) = -\vec{\mu}\cdot\vec{B}_1(t) = -\frac{1}{2}\,\hbar\omega_1(\sigma_x\cos\omega t - \sigma_y\sin\omega t) \tag{5.23}$$

où $\omega_1 = \gamma B_1$ est la *fréquence de Rabi*, souvent appelée *fréquence de nutation* $\omega_{\rm nut}$ en RMN. Il sera commode pour écrire H_1 d'utiliser les matrices $\sigma_\pm = (\sigma_x \pm {\rm i}\sigma_y)/2$ qui obéissent aux relations de commutation

$$[\sigma_z, \sigma_\pm] = \pm 2\sigma_\pm \tag{5.24}$$

H_1 prend alors la forme

$$H_1(t) = -\frac{1}{2}\,\hbar\omega_1\left[\sigma_+\,{\rm e}^{{\rm i}\omega t} + \sigma_-\,{\rm e}^{-{\rm i}\omega t}\right] \tag{5.25}$$

Rappelons que le hamiltonien *complet* dépendant du temps est

$$H(t) = H_0 + H_1(t) \tag{5.26}$$

Notre objectif est de résoudre l'équation d'évolution (4.11) avec le hamiltonien dépendant du temps (5.26), c'est-à-dire déterminer le vecteur d'état du spin 1/2 $|\varphi(t)\rangle$ en fonction du vecteur d'état à $t = 0$, $|\varphi(0)\rangle$. Il est possible de résoudre le système d'équations différentielles que l'on déduit de (4.11) pour les composantes (5.27) $c_\pm(t)$ de $|\varphi(t)\rangle$, mais il est plus élégant et plus rapide de se ramener à un hamiltonien indépendant du temps, qui a, de plus, l'avantage de se prêter à une interprétation géométrique. L'idée physique pour se ramener à un hamiltonien indépendant du temps est d'utiliser un référentiel $O\hat{x}\hat{y}z$ tournant autour de Oz avec la vitesse angulaire ω. Dans ce référentiel, le champ $\vec{B}_1$ est indépendant du temps : $\vec{B}_1$ est orienté en permanence suivant l'axe $O\hat{x}$. Soit $|\varphi(t)\rangle$ le vecteur d'état dans la base $\{|\pm\rangle\}$ correspondant au référentiel fixe

$$|\varphi(t)\rangle = c_+(t)|+\rangle + c_-(t)|-\rangle \tag{5.27}$$

Pour passer aux axes tournants, nous utilisons (3.59) qui nous dit que l'opérateur de rotation d'un angle θ autour de Oz est $\exp(-{\rm i}\theta\sigma_z/2)$. Les vecteurs de base dans le référentiel tournant sont donc, avec $\theta = -\omega t$

$$|\hat{+}(t)\rangle = {\rm e}^{{\rm i}\omega t\sigma_z/2}|+\rangle = {\rm e}^{{\rm i}\omega t/2}|+\rangle \qquad |\hat{-}(t)\rangle = {\rm e}^{{\rm i}\omega t\sigma_z/2}|-\rangle = {\rm e}^{-{\rm i}\omega t/2}|-\rangle$$

Dans le référentiel tournant, les composantes de $|\varphi(t)\rangle$ sont

$$\hat{c}_\pm(t) = \langle\hat{\pm}(t)|\varphi(t)\rangle = \langle\pm|{\rm e}^{-{\rm i}\omega t\sigma_z/2}|\varphi(t)\rangle = \langle\pm|\hat{\varphi}(t)\rangle$$

où nous avons défini le vecteur d'état $|\hat{\varphi}(t)\rangle$ dans le référentiel tournant par

$$|\hat{\varphi}(t)\rangle = {\rm e}^{-{\rm i}\omega t\sigma_z/2}|\varphi(t)\rangle \qquad |\hat{\varphi}(t=0)\rangle = |\varphi(t=0)\rangle \tag{5.28}$$

L'équation d'évolution pour $|\hat{\varphi}(t)\rangle$ devient

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\hat{\varphi}(t)\rangle = \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{-\mathrm{i}\omega t\sigma_z/2}|\varphi(t)\rangle\right) = \left(\frac{1}{2}\,\hbar\omega\sigma_z + \hat{H}(t)\right)|\hat{\varphi}(t)\rangle \tag{5.29}$$

avec

$$\hat{H}(t) = \mathrm{e}^{-\mathrm{i}\omega t\sigma_z/2}\,H(t)\,\mathrm{e}^{\,\mathrm{i}\omega t\sigma_z/2} = H_0 + \hat{H}_1(t) \tag{5.30}$$

Pour calculer $\hat{H}_1(t)$, nous allons partir de la forme (5.25) de $H_1(t)$. Les opérateurs $\hat{\sigma}_\pm(t)$

$$\hat{\sigma}_\pm(t) = \mathrm{e}^{-\mathrm{i}\omega t\sigma_z/2}\,\sigma_\pm\,\mathrm{e}^{\,\mathrm{i}\omega t\sigma_z/2} \tag{5.31}$$

obéissent à l'équation différentielle

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\sigma}_\pm(t) = -\frac{\mathrm{i}}{2}\,\omega\mathrm{e}^{-\mathrm{i}\omega t\sigma_z/2}\,[\sigma_z,\sigma_\pm]\,\mathrm{e}^{\,\mathrm{i}\omega t\sigma_z/2} = \mp\mathrm{i}\omega\hat{\sigma}_\pm(t)$$

où nous nous sommes servis de (5.24), d'où le résulat important que nous aurons souvent l'occasion d'utiliser

$$\boxed{\hat{\sigma}_\pm(t) = \mathrm{e}^{\mp\mathrm{i}\omega t}\,\sigma_\pm} \tag{5.32}$$

On en déduit immédiatement l'expression de $\hat{H}_1$, qui comme promis est indépendant du temps

$$\hat{H}_1 = -\frac{1}{2}\,\hbar\omega_1[\sigma_+ + \sigma_-] = -\frac{1}{2}\,\hbar\omega_1\sigma_x \tag{5.33}$$

La proportionnalité de $\hat{H}_1$ à σ_x ne devrait pas surprendre, puisque $\vec{B}_1$ est aligné suivant $O\hat{x}$ dans le référentiel tournant, et on aurait pu anticiper (5.33). Combinant (5.29), (5.30) et (5.33), on déduit l'expression du hamiltonien $\hat{H}$ dans le référentiel tournant

$$\boxed{\hat{H} = \frac{1}{2}\,\hbar\delta\sigma_z - \frac{1}{2}\,\hbar\omega_1\sigma_x} \tag{5.34}$$

Le *désaccord* $\delta = \omega - \omega_0$ est défini comme la différence entre la fréquence de rotation de $\vec{B}_1$ et la fréquence de Larmor.

5.2.2 Oscillations de Rabi

Comme $\hat{H}$ est indépendant du temps, l'opérateur d'évolution en axes tournants est $\hat{U}(t) = \exp(-\mathrm{i}\hat{H}t/\hbar)$. Le cas le plus simple est celui de la résonance,[6] $\delta = 0$. Dans le cas résonant, l'opérateur d'évolution se réduit à

$$\hat{U}(t) = \mathrm{e}^{\mathrm{i}\omega_1 t\sigma_x/2} = I\cos\frac{\omega_1 t}{2} + \mathrm{i}\sigma_x\sin\frac{\omega_1 t}{2} \tag{5.35}$$

6. Nous verrons ultérieurement dans cette sous-section la raison de cette terminologie.

Par comparaison avec (3.59), nous voyons que $\hat{U}(t)$ est l'opérateur de rotation d'un angle $\theta = -\omega_1 t$ autour de l'axe $O\hat{x}$. Supposons que nous partions au temps $t = 0$ de l'état $|\hat{\varphi}(0)\rangle = |+\rangle$; au temps t, compte tenu de $\sigma_x|+\rangle = |-\rangle$, le vecteur d'état est

$$|\hat{\varphi}(t)\rangle = \cos\frac{\omega_1 t}{2}\,|+\rangle + \mathrm{i}\sin\frac{\omega_1 t}{2}\,|-\rangle \tag{5.36}$$

La probabilité $\mathsf{p}_-(t)$ d'observer le spin dans l'état $|-\rangle$ est une fonction oscillante de t

$$\mathsf{p}_-(t) = |\langle -|\hat{\varphi}(t)\rangle|^2 = \sin^2\frac{\omega_1 t}{2} \tag{5.37}$$

Le spin passe donc périodiquement d'un niveau à l'autre, et ces oscillations sont appelées *oscillations de Rabi*. À nouveau, l'interprétation géométrique de (5.35) est claire : à la résonance, la précession de Larmor est de fréquence $\omega_0 = \omega$, et cette précession est exactement compensée par la rotation du référentiel. Il reste donc uniquement la précession de Larmor due à $\vec{B}_1$ qui s'effectue autour de $O\hat{x}$, puisque $\vec{B}_1$ est aligné suivant $O\hat{x}$ (figure 5.7a).

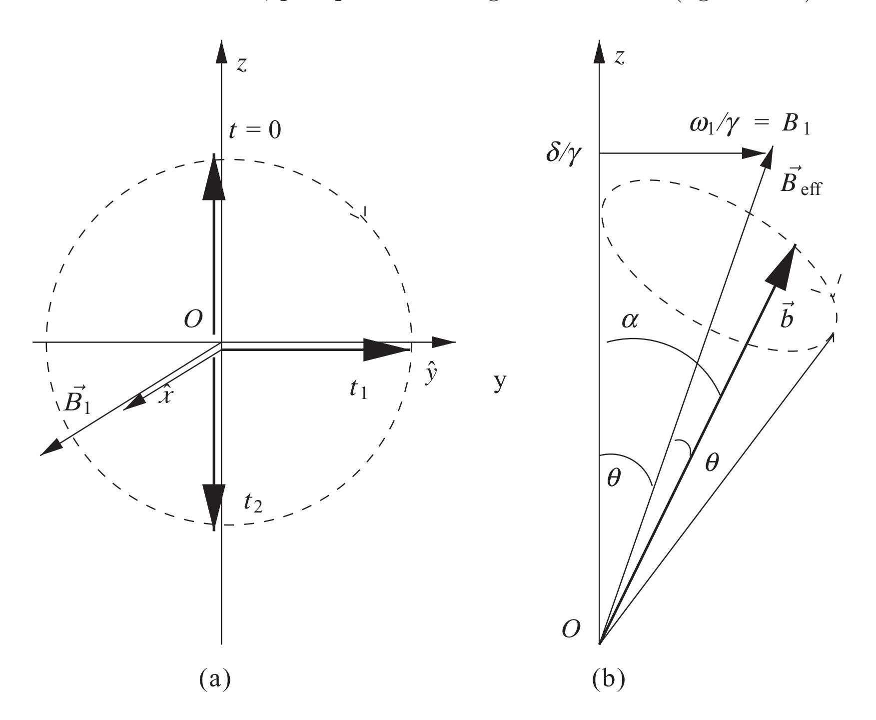


Fig. 5.7 – Schéma de l'évolution temporelle du spin dans le référentiel tournant ; le vecteur de Bloch $\vec{b} = \langle\vec{\sigma}\rangle$ est représenté par une flèche épaisse dont l'extémité se trouve sur une sphère de rayon unité. (a) À la résonance : impulsions $\pi/2$ ($t = t_1$) et π ($t = t_2$). (b) Hors résonance : on part à $t = 0$ de l'état $|+\rangle$ et l'angle θ entre $\vec{b}$ et $\vec{B}_{\text{eff}}$ est constant.

Il existe deux cas particuliers importants de (5.37). Le spin initialement dans l'état $|+\rangle$ sera trouvé dans l'état $|-\rangle$ pour des temps t donnés par

$$\frac{\omega_1 t}{2} = \left(n + \frac{1}{2}\right)\pi \qquad n = 0, 1, 2, 3, \ \ldots \tag{5.38}$$

Si le champ de radiofréquences est appliqué pendant un temps t vérifiant (5.38), en général avec $n = 0$, on dit que l'on a appliqué une *impulsion* π. Dans l'interprétation géométrique, le spin orienté initialement suivant la direction des z positifs se retrouve dans la direction des z négatifs (figure 5.7a). Lorsque

$$\frac{\omega_1 t}{2} = \left(n + \frac{1}{2}\right)\frac{\pi}{2} \qquad n = 0, 1, 2, 3, \ \ldots \tag{5.39}$$

on dit que l'on a appliqué une *impulsion* $\pi/2$. Le spin se retrouve alors dans une combinaison linéaire à poids égaux des états $|+\rangle$ et $|-\rangle$, et dans l'interprétation géométrique, le spin orienté initialement suivant Oz se retrouve suivant $O\hat{y}$ (figure 5.7a).

Hors résonance, lorsque $\delta \neq 0$, on définit un vecteur unitaire $\hat{n}$ par

$$\hat{n} \quad : \quad n_x = -\frac{\omega_1}{\Omega} \qquad n_y = 0 \qquad n_z = \frac{\delta}{\Omega} \tag{5.40}$$

où la fréquence Ω vaut

$$\Omega = \sqrt{\omega_1^2 + \delta^2} = \sqrt{\omega_1^2 + (\omega - \omega_0)^2} \tag{5.41}$$

Avec ces définitions, on peut écrire l'opérateur d'évolution $\hat{U}$ sous la forme d'un opérateur de rotation, et utiliser (3.49) pour le calculer

$$\begin{aligned}\hat{U}(t) &= \exp\left(-\mathrm{i}\frac{\Omega t}{2}\left[\sigma_x n_x + \sigma_y n_y + \sigma_z n_z\right]\right) = \exp\left(-\mathrm{i}\frac{\Omega t}{2}\,\vec{\sigma}\cdot\hat{n}\right) \\ &= I\,\cos\frac{\Omega t}{2} - \mathrm{i}(\vec{\sigma}\cdot\hat{n})\,\sin\frac{\Omega t}{2} = I\,\cos\frac{\Omega t}{2} + \mathrm{i}\left(\sigma_x\frac{\omega_1}{\Omega} - \sigma_z\frac{\delta}{\Omega}\right)\sin\frac{\Omega t}{2}\end{aligned} \tag{5.42}$$

Si l'on part au temps $t = 0$ de l'état $|+\rangle$, la probabilité d'observer le spin dans l'état $|-\rangle$ au temps t sera

$$\mathsf{p}_-(t) = |\langle -|\hat{U}(t)|+\rangle|^2 = \frac{\omega_1^2}{\Omega^2}\sin^2\frac{\Omega t}{2} \tag{5.43}$$

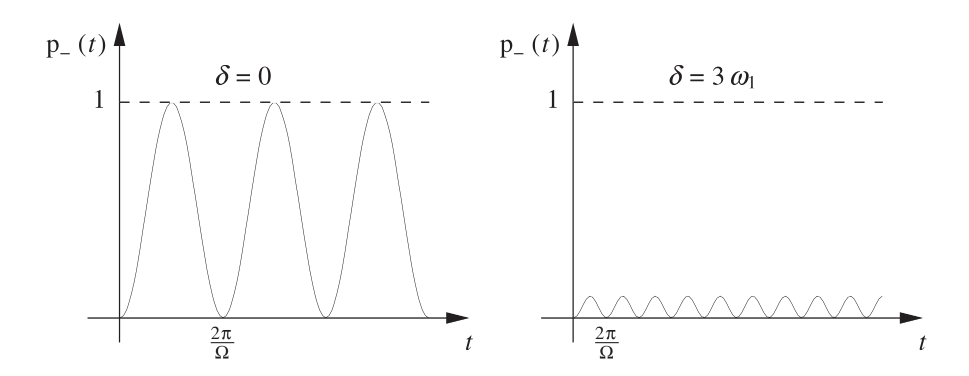


FIG. 5.8 – Oscillations de Rabi : (a) $\delta = 0$, (b) $\delta = 3\omega_1$. Dans le cas (b), la valeur maximale de $\mathsf{p}_-(t)$ est 1/10.

On voit que la probabilité maximale de transfert de l'état $|+\rangle$ vers l'état $|-\rangle$ pour $\Omega t/2 = \pi/2$ est donnée par une *courbe de résonance* de largeur δ

$$\mathsf{p}_-^{\max} = \frac{\omega_1^2}{\Omega^2} = \frac{\omega_1^2}{\omega_1^2 + \delta^2} = \frac{\omega_1^2}{\omega_1^2 + (\omega - \omega_0)^2} \tag{5.44}$$

Comme le montre la figure 5.8, les oscillations de Rabi sont maximales à la résonance, et elles diminuent rapidement d'amplitude quand δ croît. L'interprétation intuitive est claire : l'influence du champ de radiofréquences $\vec{B}_1$ est maximale lorsque celui-ci tourne à la même vitesse que le spin animé par précession de Larmor autour de $\vec{B}_0$.

Revenons à l'interprétation géométrique de ces résultats hors résonance. Si $\vec{B}_1 = 0$, la précession de Larmor s'effectue autour de Oz dans le référentiel tournant à la vitesse angulaire $\delta = \omega - \omega_0$, et si $\vec{B}_1 \neq 0$, le champ magnétique effectif $\vec{B}_{\rm eff}$ se compose d'un champ vertical de module δ/γ et d'un champ horizontal $\vec{B}_1$ de module ω_1/γ (figure 5.7b). La précession de Larmor s'effectue alors autour du champ $\vec{B}_{\rm eff}$ avec la fréquence angulaire Ω (5.41). Un petit exercice de géométrie (exercice 5.5.6) permet alors de retrouver (5.43).

5.2.3 Principes de la RMN et de l'IRM

La RMN est utilisée principalement pour déterminer la structure de molécules d'intérêt chimique ou biologique et pour l'étude de la matière condensée solide ou liquide. Une description détaillée du fonctionnement de la RMN nous entraînerait trop loin et nous ne ferons qu'effleurer le sujet. L'échantillon à étudier est plongé dans un champ uniforme $\vec{B}_0$ de quelques teslas, le champ maximum accessible aujourd'hui étant d'une vingtaine de teslas (figure 5.9). Si l'on veut caractériser une RMN, on donne plutôt la fréquence[7] de résonance

7. Voir la note 23 du chapitre 1.

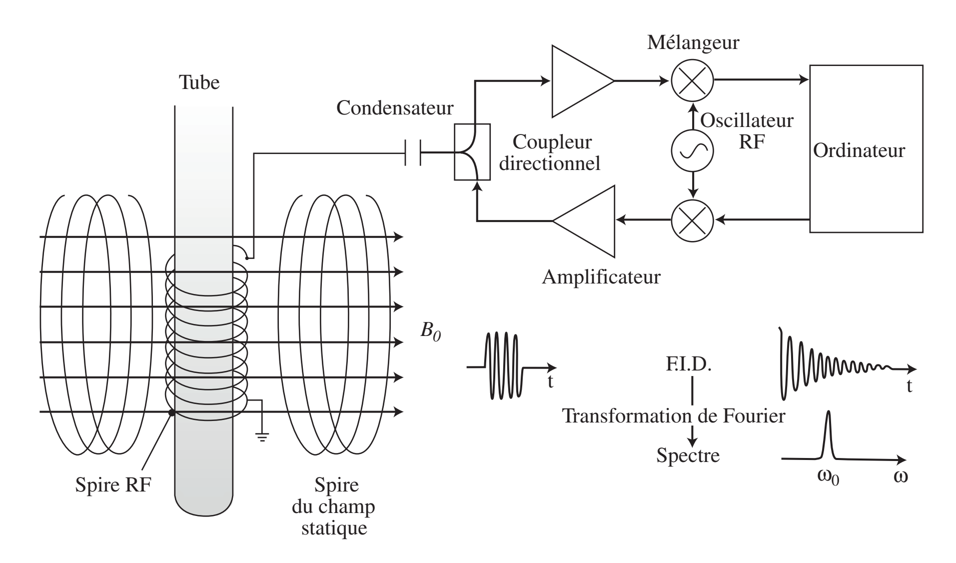


FIG. 5.9 – Schéma de principe d'une RMN. Le champ statique $\vec{B}_0$ est horizontal et le champ de radio-fréquences est généré par le solénoïde vertical. Ce solénoïde sert aussi à détecter le signal (FID = Free Induction Decay). L'impulsion RF et le signal sont dessinés en bas à gauche de la figure. On notera la décroissance exponentielle du signal et le pic de sa tranformée de Fourier à $\omega = \omega_0$. Adapté de Nielsen et Chuang [2001].

$\nu_0 = \omega_0/(2\pi) = \gamma B_0/(2\pi)$ pour un proton : un champ de 1 tesla correspond à une fréquence $\simeq 42.5$ MHz, et on parlera donc d'une RMN de 600 MHz si le champ B_0 vaut 14 teslas. En raison de la loi de Boltzmann (1.12), le niveau $|+\rangle$ est plus peuplé que le niveau $|-\rangle$, du moins si $\gamma > 0$, ce qui est le cas usuel

$$\frac{\mathsf{p}_+(t=0)}{\mathsf{p}_-(t=0)} = \exp\left(\frac{\hbar\omega_0}{k_\mathrm{B}T}\right) \tag{5.45}$$

À la température ambiante et pour une RMN de 600 MHz, la différence de population

$$\mathsf{p}_+ - \mathsf{p}_- \simeq \frac{\hbar\omega_0}{2k_\mathrm{B}T}$$

entre les niveaux $|+\rangle$ et $|-\rangle$ est $\sim 5\times 10^{-5}$.

L'application d'un champ de radiofréquences $\vec{B}_1(t)$ voisin de la résonance pendant un temps t tel que $\omega_1 t = \pi$, ou impulsion π (*cf.* (5.38)), fait passer les spins de l'état $|+\rangle$ vers l'état $|-\rangle$, provoquant donc une *inversion de population* par rapport à celle de l'équilibre, et l'échantillon est hors équilibre.

Le retour à l'équilibre est contrôlé par un temps de relaxation[8] noté T_1, le *temps de relaxation longitudinale*. On utilise en général une impulsion $\pi/2$, $\omega_1 t = \pi/2$. Ceci correspond géométriquement à une rotation du spin d'un angle $\pi/2$ autour d'un axe du plan xOy : si le spin est initialement parallèle à $\vec{B}_0$, il se retrouve dans un plan perpendiculaire à $\vec{B}_0$, un plan transversal (tandis qu'une impulsion π amène le spin dans la direction longitudinale $-\vec{B}_0$, voir la figure 5.7a). Le retour à l'équilibre est alors contrôlé par un temps de relaxation noté T_2, le *temps de relaxation transverse*. En général, $T_2 \ll T_1$, et le retour à l'équilibre est plus rapide pour une impulsion $\pi/2$ que pour un impulsion π : c'est pourquoi l'impulsion $\pi/2$ est préférée. Dans tous les cas, le retour à l'équilibre se fait en engendrant un champ magnétique tournant à la fréquence ω_0 dû au mouvement de rotation de l'ensemble des spins qui forment un dipôle macroscopique, et l'analyse de Fourier du signal donne un spectre de fréquences qui permet de remonter à la structure de la molécule étudiée. Pour ce faire, on se fonde principalement sur les propriétés suivantes :

- La fréquence de résonance dépend des noyaux par l'intermédiaire de γ.
- Pour un même noyau, la fréquence de résonance est légèrement modifiée par l'environnement chimique de l'atome correspondant, ce que l'on peut traduire en définissant un champ magnétique effectif B'_0 agissant sur le noyau
$$B'_0 = (1-\sigma)B_0 \qquad \sigma \sim 10^{-6}$$
σ est appelé le *déplacement chimique*, et il existe des corrélations fortes entre σ et la nature du groupement chimique auquel appartient le noyau considéré.
- Les interactions entre spins nucléaires voisins provoquent un clivage des fréquences de résonance en plusieurs sous-fréquences, également caractéristiques des groupements chimiques.

Ceci est résumé sur la figure 5.10 qui donne un spectre RMN typique.

Dans le cas de l'imagerie par résonance magnétique[9] (IRM), on s'intéresse exclusivement aux protons contenus dans l'eau et les graisses. L'échantillon est placé dans un champ $\vec{B}_0$ non uniforme, ce qui fait que la fréquence de résonance dépend du point d'espace. Comme l'amplitude du signal est directement proportionnelle à la densité des spins, et donc à celle des protons, on peut en déduire, après des calculs informatiques complexes, une image tridimensionnelle de la densité d'eau dans les tissus biologiques. La résolution spatiale est aujourd'hui de l'ordre du millimètre, et on peut faire une image en 0.1 s. Ceci a permis le développement de l'IMR fonctionnelle (IMRf), grâce

8. Lorsque l'on applique un champ $\vec{B}_0$, l'équilibre thermodynamique (5.45) ne s'établit pas instantanément, mais seulement après un temps $\sim T_1$.

9. L'adjectif nucléaire a été supprimé pour ne pas effrayer le public !

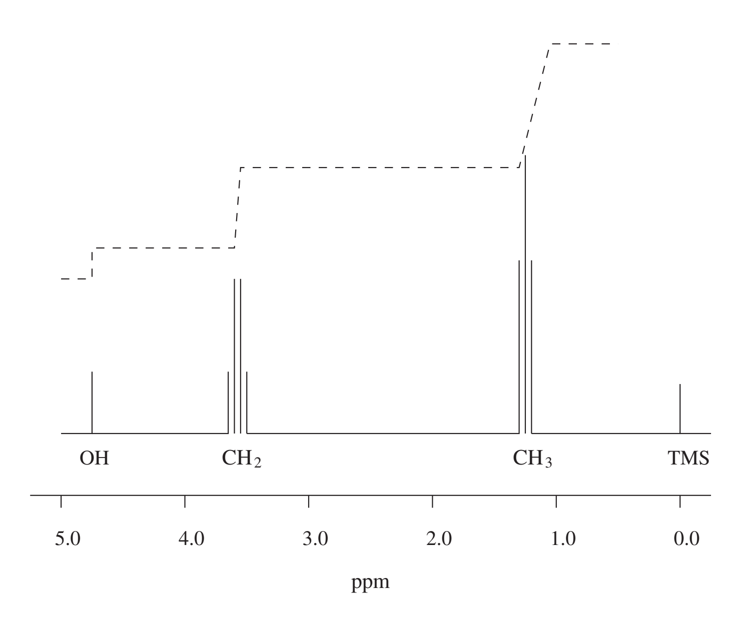


Fig. 5.10 – Spectre RMN des protons de l'éthanol CH_3CH_2OH obtenus avec une RMN de 200 MHz. On observe trois pics associés au trois groupements OH, CH_3 et CH_2. La courbe en tirets représente l'aire intégrée des signaux. Le signal TMS est un signal de référence.

à laquelle peut par exemple "voir le cerveau en action" en mesurant les variations locales de débit sanguin. Les temps de relaxation longitudinale T_1 et transverse T_2 jouent un grand rôle dans l'obtention et l'interprétation des signaux de l'IMR.

Nous allons rencontrer à nouveau les oscillations de Rabi entre deux niveaux dans les deux sections suivantes. Cependant, il existe une importante différence de principe entre la RMN et les systèmes étudiés dans ces deux sections. Nous y reviendrons à la fin du § 5.4.1.

5.3 La molécule d'ammoniac

La molécule d'ammoniac nous fournit un second exemple concret d'un système à deux niveaux que l'on peut coupler à un champ extérieur périodique.

5.3.1 La molécule d'ammoniac comme système à deux niveaux

La molécule d'ammoniac a une forme pyramidale, où l'atome d'azote occupe le sommet de la pyramide et où les trois atomes d'hydrogène forment un triangle équilatéral qui constitue la base de la pyramide (figure 5.11). Les mouvements

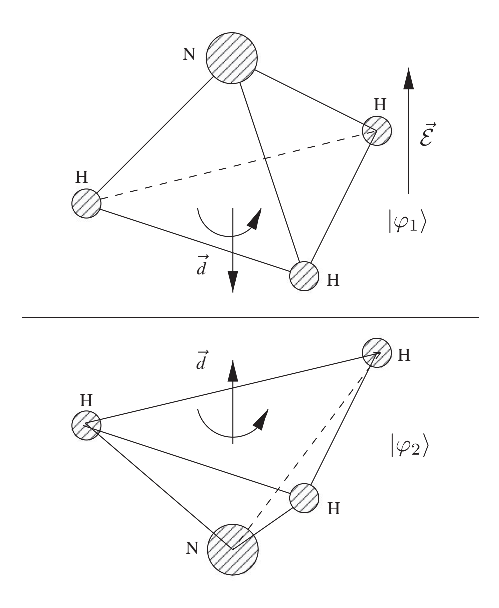
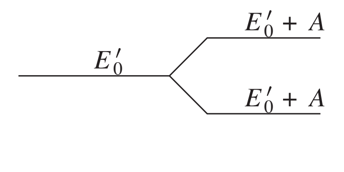


FIG. 5.11 – Les deux configurations de la molécule d'ammoniac.

possibles de cette molécule sont très variés : elle peut effectuer des mouvements de translation et de rotation dans l'espace, les atomes peuvent vibrer autour de leur position d'équilibre, les électrons peuvent se trouver dans des états excités. Une fois fixés les degrés de liberté de translation, rotation et vibration pour la molécule dans son état fondamental électronique, il reste encore deux configurations possibles pour la molécule en rotation[10] autour de son axe de symétrie, qui sont symétriques l'une de l'autre par réflexion par rapport à un plan (figure 5.11). Pour passer d'une configuration à l'autre, l'atome d'azote doit traverser le plan des atomes d'hydrogène. Ceci est possible grâce à un effet tunnel, que nous expliquerons au § 12.4.5. Dans ce qui suit, nous allons nous intéresser *uniquement* à ces deux configurations, ce qui est justifié en raison des énergies mises en jeu (*cf.* la note 11). Comme dans le cas de la molécule d'éthylène, nous utiliserons pour décrire ces deux configurations un espace des états à deux dimensions : la molécule dans l'état 1 (resp. 2) de la figure 5.11 sera décrite par le vecteur de base $|\varphi_1\rangle$ (resp. $|\varphi_2\rangle$). Si l'atome d'azote ne pouvait jamais franchir le plan des atomes d'hydrogène, l'énergie des états $|\varphi_1\rangle$ et $|\varphi_2\rangle$ serait identique, égale à E_0. Mais il existe une amplitude

10. L'importance de cette rotation pour générer deux configurations différentes est soulignée par Feynman ; dans les exposés qui ont repris ultérieurement sa présentation originale, ce mouvement de rotation a souvent été oublié. Mais si cette rotation est absente, on passe continûment d'une configuration à l'autre par une rotation dans l'espace !



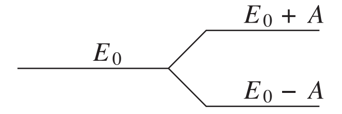


FIG. 5.12 – Clivage de deux niveaux E_0 et E_0'.

non nulle pour franchir ce plan, et le hamiltonien prend la forme (5.2)

$$H = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} \tag{5.46}$$

avec bien sûr des valeurs de E_0 et A différentes de celles de la section 5.1. La valeur de E_0 n'est pas importante pour notre discussion. En revanche, il vaut la peine d'observer que la valeur de A dans (5.46) diffère de celle de (5.2) par plusieurs ordres de grandeur. En effet, nous avons maintenant $2A \simeq 10^{-4}\,$eV, alors que précédemment $2A$ était de l'ordre de 1 eV : ceci reflète le fait qu'il est facile à un électron π de sauter d'un atome à l'autre, alors qu'il est très difficile à l'atome d'azote de franchir le plan des atomes d'hydrogène. Cette énergie de $10^{-4}\,$eV correspond à une fréquence de 24 GHz, ou à une longueur d'onde de 1.25 cm, dans le domaine des ondes centimétriques. Elle est très faible par rapport aux énergies d'excitation des électrons (quelques eV), faible par rapport aux énergies de vibration ($\sim 0.1\,$eV) et même de rotation[11]($\sim 10^{-3}\,$eV). Cette comparaison justifie l'approximation par un système à deux niveaux, car la différence entre deux niveaux de rotation successifs est de l'ordre de $10\,A$ (figure 5.12). Cependant, la molécule n'est pas dans son niveau de rotation fondamental, car $k_\mathrm{B}T \sim 0.025\,$eV est grand par rapport à $\sim 10^{-3}\,$eV : les niveaux de rotation sont excités thermiquement.

Suivant la discussion du § 5.1.1, les niveaux d'énergie de H sont $E_0 \mp A$, correspondant aux états stationnaires (5.2) et (5.3)

$$\begin{aligned} E_0 - A: &\quad |\chi_+\rangle = \frac{1}{\sqrt{2}}\Big(|\varphi_1\rangle + |\varphi_2\rangle\Big) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \\ E_0 + A: &\quad |\chi_-\rangle = \frac{1}{\sqrt{2}}\Big(|\varphi_1\rangle - |\varphi_2\rangle\Big) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} \end{aligned} \tag{5.47}$$

11. La molécule d'ammoniac possède deux fréquences propres de rotation, dont une dégénérée, correspondant à des énergies de 0.8×10^{-3} eV et de 1.2×10^{-3} eV (dégénérée), et quatre modes de vibration dont deux dégénérés, l'énergie la plus faible étant de 0.12 eV. De plus, il faudrait tenir compte des complications dues à la structure hyperfine.

L'état symétrique $|\chi_+\rangle$ est l'état fondamental, d'énergie $(E_0 - A)$ et l'état antisymétrique $|\chi_-\rangle$ est l'état excité, d'énergie $(E_0 + A)$.

5.3.2 La molécule dans un champ électrique : le maser à ammoniac

La molécule d'ammoniac possède un moment dipolaire électrique $\vec{d}$ qui, par symétrie, est perpendiculaire au plan des atomes d'hydrogène. Comme les atomes d'hydrogène ont tendance à perdre leurs électrons et l'atome d'azote à les attirer, ce moment dipolaire est orienté de l'atome d'azote vers le plan des atomes d'hydrogène (figure 5.11). Plaçons la molécule dans un champ électrique $\vec{\mathcal{E}}$ dirigé suivant Oz. L'énergie d'un dipôle classique $\vec{d}$ dans un champ électrique $\vec{\mathcal{E}}$ (nous utilisons une lettre calligraphiée pour le champ électrique, afin d'éviter toute confusion avec l'énergie) est

$$E = -\vec{d} \cdot \vec{\mathcal{E}} \tag{5.48}$$

En mécanique quantique, le moment dipolaire est un opérateur $\vec{D}$, qui s'exprime en fonction des charges et des opérateurs position des différentes particules chargées. Nous admettrons que la restriction de $\vec{D}$ à notre sous-espace à deux dimensions est donnée par la matrice suivante dans la base $\{|\varphi_1\rangle, |\varphi_2\rangle\}$

$$-\vec{D} \to \begin{pmatrix} d & 0 \\ 0 & -d \end{pmatrix} \qquad - \vec{D} \cdot \vec{\mathcal{E}} \to \begin{pmatrix} d\,\mathcal{E} & 0 \\ 0 & -d\,\mathcal{E} \end{pmatrix}$$

Ceci correspond bien au schéma de la figure 5.11 : en effet, l'énergie de l'état $|\varphi_1\rangle$ de la figure 5.11 est $+d\,\mathcal{E}$ car le moment dipolaire est antiparallèle au champ, et celle de l'état $|\varphi_2\rangle$ est $-d\,\mathcal{E}$ car le moment dipolaire est parallèle au champ. En dernier ressort, la forme matricielle de ce moment dipolaire est justifiée par l'accord avec l'expérience. Le hamiltonien prend donc la forme

$$H = \begin{pmatrix} E_0 + d\,\mathcal{E} & -A \\ -A & E_0 - d\,\mathcal{E} \end{pmatrix} \tag{5.49}$$

Nous examinons d'abord le cas d'un champ électrique *statique*. Le hamiltonien est alors indépendant du temps. Le calcul des valeurs propres est immédiat[12]

$$\det \begin{pmatrix} E_0 + d\,\mathcal{E} - E & -A \\ -A & E_0 - d\,\mathcal{E} - E \end{pmatrix} = (E - E_0)^2 - (d\,\mathcal{E})^2 - A^2 = 0$$

soit

$$E_\pm = E_0 \mp \sqrt{A^2 + (d\mathcal{E})^2} \tag{5.50}$$

Ces valeurs propres sont représentées sur la figure 5.13 en fonction de $\mathcal{E}$. Si $d\,\mathcal{E} \gg A$, les énergies sont $\simeq E_0 \pm d\mathcal{E}$ et les vecteurs propres correspondants

12. On peut aussi utiliser les résultats du § 2.3.2.

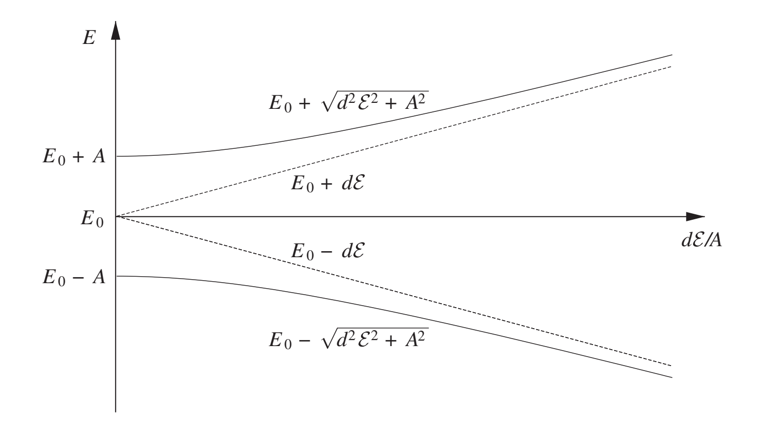


FIG. 5.13 – Valeurs de l'énergie en fonction du champ électrique $\mathcal{E}$.

approximativement $|\varphi_1\rangle$ et $|\varphi_2\rangle$. En pratique, on se trouve dans le cas opposé : $d\,\mathcal{E} \ll A$. On peut alors développer la racine carrée dans (5.50)

$$E_\pm \simeq E_0 \mp A \mp \frac{1}{2}\,\frac{d^2\mathcal{E}^2}{A} \tag{5.51}$$

À des termes d'ordre $d\,\mathcal{E}/(2A)$ près (exercice 5.5.4), les vecteurs propres sont $|\chi_+\rangle$ et $|\chi_-\rangle$. Si le champ électrique n'est pas uniforme, la molécule sera soumise à une force

$$\vec{F}_\pm = -\vec{\nabla} E_\pm = \pm\frac{d^2}{2A}\,\vec{\nabla}\mathcal{E}^2 \tag{5.52}$$

Comme dans l'expérience de Stern-Gerlach, on pourra séparer expérimentalement les états propres $|\chi_\pm\rangle$ du hamiltonien (5.18) en utilisant un champ électrique inhomogène[13] : voir la figure 5.15.

Le schéma de niveaux que nous venons de trouver est très général : il met en évidence le phénomène de *répulsion des niveaux* (figure 5.14). Si $d\mathcal{E} \gg A$, les états propres du hamiltonien sont $|\varphi_1\rangle$ et $|\varphi_2\rangle$. Pour $\mathcal{E} = 0$, ces deux niveaux devraient se croiser et échanger leur stabilité. Il n'en est rien, à cause de la valeur non nulle de A, qui entraîne que les deux niveaux ne se croisent pas.

Supposons maintenant que le champ électrique est un champ *oscillant*

$$\mathcal{E}(t) = \mathcal{E}_0 \cos\omega t = \frac{1}{2}\,\mathcal{E}_0\Big(\mathrm{e}^{\,\mathrm{i}\omega t} + \mathrm{e}^{-\mathrm{i}\omega t}\Big) \qquad \mathcal{E}_0 \ \text{ réel} > 0 \tag{5.53}$$

Le hamiltonien dépend explicitement du temps. Il sera commode de prendre comme vecteurs de base les états stationnaires (5.47) $|\chi_+\rangle$ et $|\chi_-\rangle$ du

13. En pratique, le champ est choisi tel que l'état $|\chi_-\rangle$ soit focalisé et l'état $|\chi_+\rangle$ défocalisé : *cf.* Basdevant et Dalibard [2001], chapitre 6.

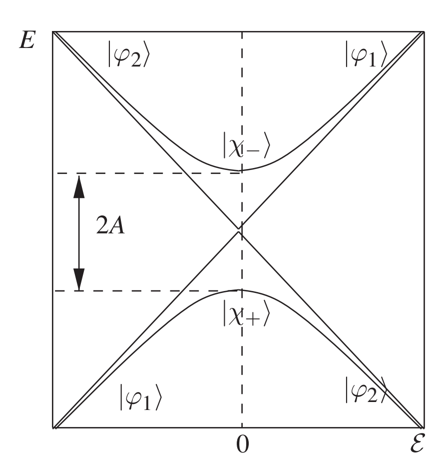


Fig. 5.14 – Le phénomène de répulsion des niveaux.

hamiltonien (5.46), plutôt que $|\varphi_1\rangle$ et $|\varphi_2\rangle$. Le hamiltonien (5.49) devient dans cette nouvelle base

$$H(t) = \begin{pmatrix} E_0 - A & d\,\mathcal{E}(t) \\ d\,\mathcal{E}(t) & E_0 + A \end{pmatrix} = E_0\, I - A\,\sigma_z + d\mathcal{E}_0\,\sigma_x\,\cos\omega t = E_0 I + H_0 + W(t) \tag{5.54}$$

Ce hamiltonien est manifestement très semblable à celui du § 5.2.1, et nous pourrions utiliser les mêmes techniques pour résoudre l'équation d'évolution (4.11). Toutefois, afin d'illustrer l'utilisation du point de vue de l'interaction, nous allons plutôt nous placer dans ce cadre en prenant pour hamiltonien H_0 (*cf.* le § 4.2.5, le terme $E_0\, I$ ne joue aucun rôle)

$$H_0 = -A\sigma_z = -\frac{1}{2}\,\hbar\omega_0\,\sigma_z \tag{5.55}$$

Nous avons posé $\omega_0 = 2A/\hbar$, qui représente physiquement la fréquence angulaire $\simeq 1.5 \times 10^{12}\,\text{rad.s}^{-1}$ d'une onde électromagnétique émise lorsque la molécule passe du niveau excité d'énergie $(E_0 + A)$ au niveau fondamental d'énergie $(E_0 - A)$: $2A$ est l'énergie du photon émis dans cette transition. La fréquence ω_0 est à nouveau appelée fréquence de résonance. Suivant (4.40), nous appelons $|\tilde{\varphi}(t)\rangle$ le vecteur d'état dans le point de vue de l'interaction

$$|\tilde{\varphi}(t)\rangle = \mathrm{e}^{\,\mathrm{i}H_0 t/\hbar}\,|\varphi(t)\rangle = \mathrm{e}^{-\mathrm{i}\hbar\omega_0\sigma_z t/2}\,|\varphi(t)\rangle \qquad |\tilde{\varphi}(t=0)\rangle = |\varphi(t=0)\rangle \tag{5.56}$$

Comparant avec (5.28), nous voyons que nous avons choisi un référentiel tournant dans le sens rétrograde avec la vitesse angulaire ω_0, et non ω. Les deux référentiels tournants coïncident à la résonance. Pour transformer le terme d'interaction W, nous récrivons d'abord

$$W(t) = d\mathcal{E}_0(\sigma_+ + \sigma_-)\cos\omega t = \frac{1}{2}\,d\mathcal{E}_0\left[\sigma_+\,\mathrm{e}^{\,\mathrm{i}\omega t} + \sigma_+\,\mathrm{e}^{-\mathrm{i}\omega t} + \sigma_-\,\mathrm{e}^{\,\mathrm{i}\omega t} + \sigma_-\,\mathrm{e}^{-\mathrm{i}\omega t}\right]$$

et nous utilisons l'analogue de (5.32)

$$\tilde{\sigma}_{\pm}(t) = \mathrm{e}^{\mp \mathrm{i}\omega_0 t}\, \sigma_{\pm} \tag{5.57}$$

pour obtenir

$$\tilde{W}(t) = \frac{1}{2}\, d\mathcal{E}_0 \left[\sigma_+\, \mathrm{e}^{\mathrm{i}\delta t} + \sigma_-\, \mathrm{e}^{-\mathrm{i}\delta t} + \sigma_+\, \mathrm{e}^{-\mathrm{i}(\omega+\omega_0)t} + \sigma_-\, \mathrm{e}^{\mathrm{i}(\omega+\omega_0)t} \right] \tag{5.58}$$

À la résonance, les deux premiers termes du crochet redonnent le résultat (5.34) de la RMN, mais deux termes supplémentaires ont fait leur apparition. En fait, ceci est dû au choix d'une polarisation linéaire pour $\vec{\mathcal{E}}$, car dans le cas d'une polarisation circulaire, nous serions retombés sur les équations du § 5.2.1. Dans le cas de la RMN, nous aurions également vu apparaître ces termes supplémentaires si nous avions choisi un champ de radiofréquences $\vec{B}_1$ orienté suivant un axe fixe (exercice 5.5.6)

$$\vec{B}_1(t) = 2\hat{x}\, B_1 \cos \omega t$$

Ces termes supplémentaires peuvent être négligés si deux conditions sont réalisées.

- La perturbation apportée par le champ électrique est faible : $d\,\mathcal{E}_0 \ll A$, ou de façon équivalente, $d\,\mathcal{E}_0/\hbar \ll \omega_0$. La fréquence de Rabi est cette fois la quantité $\omega_1 = d\,\mathcal{E}_0/\hbar$. La condition de champ faible est donc aussi $\omega_1 \ll \omega_0$, ce qui est – presque – toujours réalisé en pratique.
- La deuxième hypothèse est que la fréquence du champ électrique soit proche de la résonance : $\omega \simeq \omega_0$, condition qui s'exprime en fonction du désaccord $\delta = (\omega - \omega_0)$ et qui s'écrit plus précisément $|\delta| \ll \omega_0$. Dans ces conditions, les termes en $\exp[\pm \mathrm{i}(\omega + \omega_0)t]$ de (5.58) varient très rapidement par rapport aux termes en $\exp(\pm \mathrm{i}\delta t)$ et leur effet moyenné dans le temps est négligeable.

Si ces deux conditions sont vérifiées, on peut alors négliger les termes en $\exp[\pm \mathrm{i}(\omega+\omega_0)t]$ dans (5.58). Cette approximation est appelée *approximation séculaire*[14], et le hamiltonien dans le point de vue de l'interaction est

$$\boxed{\tilde{W}(t) = \frac{1}{2}\, \hbar\omega_1 \left[\sigma_+\, \mathrm{e}^{\mathrm{i}\delta t} + \sigma_-\, \mathrm{e}^{-\mathrm{i}\delta t} \right]} \tag{5.59}$$

$\tilde{W}(t)$ est indépendant du temps à la résonance

$$\tilde{W}_{\mathrm{res}}(t) = \tilde{W}_{\mathrm{res}} = \frac{1}{2}\, \hbar\omega_1 \sigma_x \tag{5.60}$$

Plaçons-nous exactement à la résonance en prenant la fréquence du champ électrique égale à la fréquence de la transition : $\omega = \omega_0$. Supposons par

14. Ou approximation des ondes tournantes, ou encore approximation quasi-résonante.

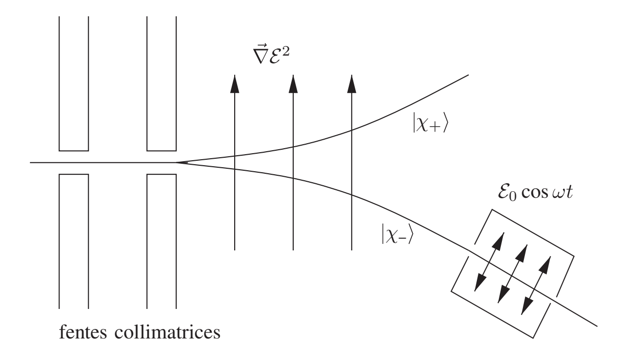


FIG. 5.15 – Maser à ammoniac.

exemple, qu'au temps $t = 0$, la molécule se trouve dans l'état $|\chi_-\rangle$[15] d'énergie $(E_0 + A)$. Pour calculer la probabilité $\mathsf{p}_\pm$ de trouver au temps t la molécule dans l'état $|\chi_\pm\rangle$, il suffit de transposer (5.37)

$$\begin{aligned} \mathsf{p}_-(t) &= |\langle\chi_-|\varphi(t)\rangle|^2 = |c_-(t)|^2 = \cos^2\left(\frac{\omega_1 t}{2}\right) \\ \mathsf{p}_+(t) &= |\langle\chi_+|\varphi(t)\rangle|^2 = |c_+(t)|^2 = \sin^2\left(\frac{\omega_1 t}{2}\right) \end{aligned} \tag{5.61}$$

La molécule passe de l'état $|\chi_-\rangle$ à l'état $|\chi_+\rangle$ avec une fréquence angulaire $\omega_1/2 = d\mathcal{E}_0/(2\hbar)$.

Après avoir mis la molécule dans l'état $|\chi_-\rangle$ grâce au filtrage décrit dans la sous-section précédente, on la fait passer dans une cavité où règne un champ oscillant à la fréquence de résonance (figure 5.15). La molécule franchit la cavité en un temps T ; si ce temps est ajusté de sorte que

$$\frac{d\,\mathcal{E}_0 T}{2\hbar} = \frac{\pi}{2}$$

à la sortie de la cavité toutes les molécules sont passées dans l'état $|\chi_+\rangle$. Par conservation de l'énergie, les molécules *fournissent* de l'énergie au champ électromagnétique : ce processus est appelé *émission stimulée (ou induite).* Si les molécules s'étaient trouvées dans l'état $|\chi_+\rangle$, elles auraient absorbé de l'énergie en l'empruntant au champ électromagnétique pour passer dans l'état $|\chi_-\rangle$, processus appelé *absorption stimulée.*

Le processus d'émission stimulée est un processus susceptible d'amplifier un champ électromagnétique, pourvu que l'on soit capable de produire les molécules dans un état excité, c'est-à-dire d'obtenir une *inversion de po-*

15. Dans le cas de la RMN, on fait passer le spin de l'état d'énergie la plus basse vers celui d'énergie la plus haute, alors que pour l'application au maser, nous nous intéressons à la situation inverse.

pulation[16]. Le dispositif expérimental représenté schématiquement sur la figure 5.15 réalise cette amplification : les molécules sélectionnées dans l'état $|\chi_-\rangle$ traversent une cavité où règne un champ électrique oscillant à la fréquence de résonance et de longueur convenablement ajustée. Ce dispositif est un prototype de maser[17].

5.3.3 Transitions hors résonance

Nous nous plaçons maintenant hors résonance : $\omega \simeq \omega_0$ mais $\omega \neq \omega_0$, et nous partons par exemple au temps $t = 0$ d'une molécule dans l'état $|\chi_+\rangle$. Nous souhaitons calculer la probabilité $\mathsf{p}(\omega; t)$ de trouver la molécule dans l'état $|\chi_-\rangle$ au temps t. À l'approximation séculaire, la résolution des

équations d'évolution donne le résultat (5.43) que nous écrivons sous la forme développée[18]

$$\mathsf{p}(\omega; t) = \frac{\omega_1^2}{(\omega - \omega_0)^2 + \omega_1^2} \sin^2\left(\frac{t}{2}\sqrt{(\omega - \omega_0)^2 + \omega_1^2}\right) \tag{5.62}$$

Rappelons que la fréquence de Rabi $\omega_1 = d\,\mathcal{E}_0/\hbar$. Bien que nous ayons été capables d'écrire une solution exacte, il est utile de donner une solution approchée simple des équations d'évolution lorsque la condition

$$\frac{d\,\mathcal{E}_0 t}{\hbar} \ll 1 \qquad t \ll \frac{\hbar}{d\mathcal{E}_0} = \tau_2 \tag{5.63}$$

est satisfaite, c'est-à-dire *pour des temps suffisamment courts*. L'intérêt de cette solution approchée est qu'elle se retrouve dans de nombreux problèmes qui ne peuvent pas être résolus exactement et elle prépare le terrain pour le chapitre 8. Nous avons à $t = 0$, pour les composantes $\tilde{c}_\pm(t)$ de $|\tilde{\varphi}(t)\rangle$

$$\tilde{c}_+ = 1 \qquad \tilde{c}_- = 0$$

Nous nous intéressons à un processus où l'absorption de rayonnement électromagnétique permet de passer du niveau fondamental au niveau excité. Dans

16. Comme nous l'avons déjà vu en (5.45), si E_0 est l'énergie de l'état fondamental et E_1 celle de l'état excité, le rapport des probabilités $\mathsf{p}_1/\mathsf{p}_0$ de trouver un système atomique ou moléculaire dans un état E_1 ou E_0 est donné par la loi de Boltzmann : $\mathsf{p}_1/\mathsf{p}_0 = \exp[(E_0 - E_1)/k_\mathrm{B}T] < 1$. Il faut donc aller à l'encontre de l'équilibre thermique pour obtenir une telle inversion de population.

17. Maser est un acronyme pour “Microwave Amplification by Stimulated Emission of Radiation” et laser pour “Light Amplification by Stimulated Emission of Radiation”.

18. Il est facile de voir que les équations d'évolution pour les composantes $\tilde{c}_\pm(t)$ dans le point de vue de l'interaction sont, à l'approximation séculaire,

$$\mathrm{i}\frac{\mathrm{d}\tilde{c}_\pm(t)}{\mathrm{d}t} = \frac{1}{2}\omega_1\,\tilde{c}_\mp(t)\,\mathrm{e}^{\pm\mathrm{i}\delta t}$$

la résolution des équations d'évolution, nous pouvons supposer $\tilde{c}_+ \simeq 1$: en effet, en raison de la condition (5.63), $\tilde{c}_+$ n'a pas le temps de varier de façon appréciable. La solution approximative de l'équation donnant $\tilde{c}_-$ est alors évidente

$$\tilde{c}_-(t) \simeq \frac{\omega_1}{2\mathrm{i}} \int_0^t \mathrm{d}t' \exp[-\mathrm{i}(\omega-\omega_0)t'] = -\frac{\omega_1}{2}\left[\frac{1-\exp[-\mathrm{i}(\omega-\omega_0)t]}{\omega-\omega_0}\right] \quad (5.64)$$

ce qui donne pour la probabilité de transition à une fréquence ω, $\mathsf{p}(\omega; t)$

$$\mathsf{p}(\omega; t) = |\tilde{c}_-(t)|^2 = \frac{1}{4}\,\omega_1^2\, t^2 \frac{\sin^2[(\omega-\omega_0)t/2]}{[(\omega-\omega_0)t/2]^2} \quad (5.65)$$

Il semble donc que $\mathsf{p}(\omega; t) \propto t^2$ pour $|\delta|t \ll 1$, mais ceci est dû au fait que le spectre de fréquences de l'onde électromagnétique contient une fréquence unique ω, ce qui est irréaliste. En réalité, on a toujours un spectre continu de fréquences, et nous allons en tenir compte. Le rapport entre ce résultat et celui obtenu à la résonance est

$$\frac{\mathsf{p}(\omega; t)}{\mathsf{p}(\omega_0; t)} = f(\omega-\omega_0; t) = \frac{\sin^2[(\omega-\omega_0)t/2]}{[(\omega-\omega_0)t/2]^2}$$

La fonction $f(\omega-\omega_0; t)$ est tracée sur la figure 5.16 en fonction de ω. Elle présente un pic aigu à $\omega = \omega_0$, de largeur $\sim 2\pi/t$. Compte tenu de

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2}\,\mathrm{d}x = \pi$$

l'aire sous la courbe est $2\pi/t$ et $f(\omega-\omega_0; t)$ est approximativement un delta de Dirac

$$f(\omega-\omega_0; t) = \frac{\sin^2[(\omega-\omega_0)t/2]}{[(\omega-\omega_0)t/2]^2} \simeq \frac{2\pi}{t}\,\delta(\omega-\omega_0) \quad (5.66)$$

Ces résultats nous permettent de calculer le taux de transition de l'état $|\chi_+\rangle$ vers l'état $|\chi_-\rangle$ dû à l'absorption de rayonnement électromagnétique par la molécule dans son état fondamental[19]. Le flux d'énergie incident $\mathcal{I}$ d'une onde électromagnétique est donné par le vecteur de Poynting $\vec{\mathcal{S}} = \varepsilon_0 c^2 \vec{\mathcal{E}} \times \vec{\mathcal{B}}$

$$\mathcal{I} = \varepsilon_0 c^2 \langle \vec{\mathcal{E}} \times \vec{\mathcal{B}} \rangle = \frac{1}{2}\,\varepsilon_0 c \mathcal{E}_0^2 \quad (5.67)$$

où $\langle \bullet \rangle$ représente une moyenne temporelle et le champ électrique est de la forme (5.53). Dans ces conditions,

$$\mathsf{p}(\omega; t) = \left(\frac{d\,\mathcal{E}_0}{2\hbar}\right)^2 t^2 f(\omega-\omega_0; t) = 2\pi\left(\frac{d^2}{4\pi\varepsilon_0\hbar^2 c}\right)\mathcal{I}\, t^2 f(\omega-\omega_0; t) \quad (5.68)$$

19. Plus précisément, il s'agit de l'ensemble des transitions d'énergie $(E_0 - A)$ à $(E_0 + A)$ (figure 5.12), ce qui suppose de sélectionner les molécules dans l'état $(E_0 - A)$ par le mécanisme décrit au § 5.2.2.

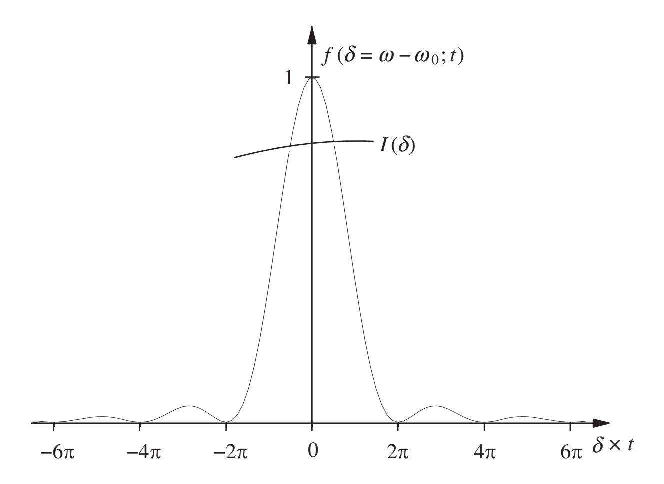


FIG. 5.16 – La fonction $f(\omega - \omega_0; t)$ définie en (5.66).

Ainsi que nous l'avons déjà signalé, la fréquence du champ électrique n'est pas exactement fixée, mais s'étale sur un spectre de fréquences $\Delta\omega$. Soit $\mathcal{I}(\omega)$ l'intensité par unité de fréquence et supposons que $\Delta\omega \gg \pi/t$ (figure 5.16) : la probabilité de transition intégrée sur ω est alors

$$\begin{aligned} \mathsf{p}(t) &= 2\pi \left(\frac{d^2}{4\pi\varepsilon_0\hbar^2 c}\right) t^2 \int_0^\infty \mathrm{d}\omega\, \mathcal{I}(\omega)\, f(\omega-\omega_0;t) \\ &\simeq 4\pi^2 \left(\frac{d^2}{4\pi\varepsilon_0\hbar^2 c}\right) \mathcal{I}(\omega_0)\, t \end{aligned}$$

où nous avons utilisé l'approximation (5.66) pour $f(\omega - \omega_0; t)$. Le fait remarquable est que $\mathsf{p}(t)$ est proportionnel à t (et non à t^2 !), et $\mathsf{p}(t)/t$ peut s'interpréter comme une *probabilité de transition par unité de temps* Γ

$$\boxed{\Gamma = \frac{1}{t}\,\mathsf{p}(t) = 4\pi^2 \left(\frac{d^2}{4\pi\varepsilon_0\hbar^2 c}\right) \mathcal{I}(\omega_0)} \tag{5.69}$$

La proportionnalité de la probabilité de transition à d^2 et à $\mathcal{I}$ est une caractéristique de la plupart des processus d'absorption du rayonnement électromagnétique par un système atomique ou moléculaire. Les conditions de validité de l'approximation sont (i) $t \gg \tau_1 \sim 1/\Delta\omega$ et (ii) $\mathsf{p}(t) \ll 1$ c'est-à-dire $t \ll \tau_2$ (*cf.* (5.63)). Il faut donc encadrer le temps t par

$$\boxed{\tau_1 \sim \frac{1}{\Delta\omega} \ll t \ll \tau_2 \sim \frac{\hbar}{d\mathcal{E}_0} = \frac{1}{\omega_1}} \tag{5.70}$$

Cette inégalité implique $\omega_1 \ll \Delta\omega$.

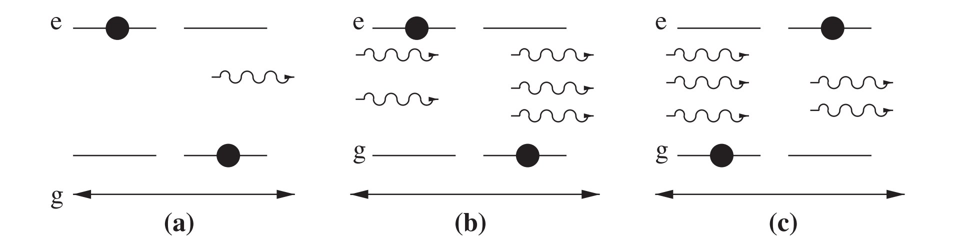


FIG. 5.17 – (a) Émission spontanée. (b) Émission stimulée. (c) Absorption.

5.4 Atome à deux niveaux

5.4.1 Absorption et émission de photons

Le calcul que nous venons de présenter a jeté les bases d'une théorie générale de l'absorption et de l'émission de rayonnement électromagnétique par un système atomique ou moléculaire, aux restrictions suivantes près.

- L'approximation par un système à deux niveaux doit être justifiée : ce sera le cas si l'on s'intéresse uniquement à des transitions entre deux niveaux séparés par une énergie $\hbar\omega_0$ et à un champ électromagnétique de fréquence $\omega \simeq \omega_0$, c'est-à-dire au voisinage de la résonance. Par convention, un des deux états, celui de plus basse énergie, sera noté $|g\rangle$ (il s'agit souvent de l'état fondamental), et le second sera noté $|e\rangle$ (état excité : figure 5.17). Dans le cas d'un atome, cette approximation est appelée *approximation de l'atome à deux niveaux*, qui fournit un modèle de base pour la physique atomique et les lasers.
- La transition doit être du type *dipolaire électrique*, c'est-à-dire contrôlée par l'élément de matrice de l'opérateur moment dipolaire électrique $\vec{D}$ entre les deux niveaux, et la condition $\omega_1 \ll \omega_0$ doit être vérifiée.
- Le champ électromagnétique est considéré comme un champ classique. Le traitement que nous venons de donner est appelé "semi-classique" : l'atome est traité comme un système quantique, mais le champ reste classique. L'aspect "photon" du champ électromagnétique est donc ignoré, et on ne peut pas en principe rendre compte de l'émission spontanée de rayonnement par un atome dans un état excité : on peut espérer au mieux en donner un traitement heuristique.
- Les résultats du § 5.3.3 doivent être modifiés pour tenir compte de la durée de vie finie de l'état excité (section 15.3).

Lorsqu'un atome à deux niveaux interagit avec un champ électromagnétique, en pratique aujourd'hui le champ d'un laser, la probabilité d'absorption se calcule selon le schéma du § 5.3.3, mais les ordres de grandeur sont tout à fait différents de ceux de la molécule d'ammoniac. En reprenant un exemple déjà mentionné au § 1.5.3, la différence d'énergie $\hbar\omega_0$ entre l'état fondamental et le premier niveau excité du rubidium est de 1.6 eV, correspondant à une

longueur d'onde de 0.78 μm, à la limite de l'infra rouge. Cet ordre de grandeur est typique de la physique atomique : les transitions généralement utilisées sont dans le domaine visible, ou bien dans le proche ultraviolet ou le proche infrarouge.

Nous avons déjà souligné que l'émission spontanée n'était pas en principe décrite par le traitement semi-classique, puisque l'on passe d'un état initial à zéro photon à un état final à un photon : un photon est *créé* au moment de la désexcitation de l'atome. Seule une théorie quantique du champ électromagnétique permet de décrire l'émission spontanée de façon rigoureuse. Bien que notre traitement classique du champ électromagnétique ne nous autorise pas une interprétation en termes de photons, nous nous risquerons néanmoins à décrire les processus du § 5.3.3 en utilisant ce concept : par exemple nous interpréterons le gain d'énergie du champ comme une augmentation du nombre de photons dans la cavité. Le processus

$$|\chi_-\rangle + n \text{ photons} \to |\chi_+\rangle + \ (n+1) \text{ photons} \tag{5.71}$$

représente donc l'émission stimulée. L'absorption stimulée est le processus inverse de (5.45)

$$|\chi_+\rangle + \ n \text{ photons} \to |\chi_-\rangle + \ (n-1) \text{ photons} \tag{5.72}$$

Enfin, l'émission spontanée d'un photon se produit quand le niveau excité $|\chi_-\rangle$ se désexcite en l'absence de champ électromagnétique

$$|\chi_-\rangle + 0 \text{ photon} \to |\chi_+\rangle + 1 \text{ photon} \tag{5.73}$$

Ces processus sont représentés schématiquement sur la figure 5.17. Il faut bien faire la différence entre l'émission stimulée, qui est cohérente avec l'onde incidente et est proportionnelle à l'intensité incidente, et l'émission spontanée qui est aléatoire, sans relation de phase avec le champ appliqué et n'est pas influencée par les conditions externes[20].

La nécessité de l'émission spontanée a été démontrée pour la première fois par Einstein. Examinons une collection d'atomes à deux niveaux E_1 et E_2, $E_1 < E_2$, placés dans une cavité à la température T. Il règne dans cette cavité un rayonnement donné par la loi de Planck (1.22). Si N est le nombre total d'atomes et $N_1(t)$, $N_2(t)$ le nombre d'atomes dans les états E_1 et E_2

$$N_1(t) + N_2(t) = N = \text{cste}$$

en supposant que seuls les états E_1 et E_2 sont peuplés de façon appréciable[21]. Les nombres $N_1(t)$ et $N_2(t)$ vérifient les équations cinétiques

$$\frac{\mathrm{d}N_1}{\mathrm{d}t} = -\frac{\mathrm{d}N_2}{\mathrm{d}t} = (-AN_1 + BN_2)\epsilon(\omega) \tag{5.74}$$

20. Sauf cas exceptionnel : si l'atome est piégé entre des miroirs hautement réfléchissants et à très basse température, il est possible de modifier l'émission spontanée. C'est ce que l'on appelle l'électrodynamique en cavité : voir le § 17.2.3, et pour plus de détails Grynberg *et al.* [2010] ou Haroche et Raimond [2006], chapitre 3.

21. Ce sera le cas si, par exemple, les autres états E_n sont tels que $E_n - E_1 \gg E_2 - E_1$ et $E_n - E_1 \gg k_\mathrm{B}T$.

où $\hbar\omega = E_2 - E_1$; $A\,\epsilon(\omega)$ est le taux de transition $E_1 \to E_2$ par unité de temps dû à l'absorption stimulée dans l'état E_1 et $B\,\epsilon(\omega)$ le taux de transition $E_2 \to E_1$ par unité de temps dû à l'émission stimulée. Ces taux sont proportionnels à la densité d'énergie $\epsilon(\omega)$. À l'équilibre,

$$\frac{\mathrm{d}N_1}{\mathrm{d}t} = \frac{\mathrm{d}N_2}{\mathrm{d}t} = 0$$

et le rapport des populations est donné par la loi de Boltzmann (1.12) d'où

$$\frac{A}{B} = \frac{N_1^{\mathrm{eq}}}{N_2^{\mathrm{eq}}} = \exp\left(-\frac{E_1 - E_2}{k_{\mathrm{B}}T}\right) = \exp\left(\frac{\hbar\omega}{k_{\mathrm{B}}T}\right) \tag{5.75}$$

Ce résultat n'est pas physiquement acceptable, car A et B ne peuvent dépendre que des caractéristiques de l'interaction du champ électromagnétique avec l'atome, et non de la température. Il faut corriger (5.75) pour tenir compte de l'émission spontanée, *indépendante* de $\epsilon(\omega)$

$$\frac{\mathrm{d}N_1}{\mathrm{d}t} = (-AN_1 + BN_2)\epsilon(\omega) + B'N_2 \tag{5.76}$$

La condition $\mathrm{d}N_1/\mathrm{d}t = 0$ jointe à la condition d'équilibre de Boltzmann donne pour $\epsilon(\omega)$

$$\epsilon(\omega) = \frac{B'}{AN_1/N_2 - B} = \frac{B'}{A\exp\left(\dfrac{\hbar\omega}{k_{\mathrm{B}}T}\right) - B} \tag{5.77}$$

La comparaison avec (1.22) montre que $A = B$ et que

$$\frac{B'}{A} = \frac{\hbar\omega^3}{\pi^2 c^3}$$

On remarque que l'on aurait aussi bien pu utiliser dans le raisonnement la densité de photons $n(\omega) = \epsilon(\omega)/\hbar\omega$ ou toute quantité proportionnelle à la densité d'énergie $\epsilon(\omega)$, au prix d'une simple redéfinition de A et B. Calculons explicitement B' ; $\epsilon(\omega)$ est une densité d'énergie par unité de fréquence, et l'intensité $\mathcal{I}(\omega)$ dans (5.67) est reliée à $\epsilon(\omega)$ par

$$\mathcal{I}(\omega) = c\,\epsilon(\omega)$$

ce qui donne par comparaison avec (5.69) la probabilité d'absorption par unité de temps

$$A = 4\pi^2 c\left(\frac{d^2}{4\pi\varepsilon_0\hbar^2 c}\right)$$

Cependant, dans notre modèle, le dipôle ne peut osciller que dans une seule direction, par exemple Oz, ce qui entraîne que les photons polarisés perpendiculairement à Oz ne sont pas absorbés, et la probabilité d'absorption de l'équation ci-dessus doit être divisée par un facteur 2. De plus, pour obtenir la

probabilité d'émission spontanée B', il faut prendre en compte une moyenne angulaire. Comme le dipôle oscille dans la direction fixe Oz le produit scalaire $\vec{d}\cdot\vec{\mathcal{E}}$ fait intervenir la polarisation $\vec{e}_s(\vec{k})$ du photon émis de vecteur d'onde $\vec{k}$, $\vec{d}\cdot\vec{\mathcal{E}} = d\hat{z}\cdot\vec{e}_s(\vec{k})$; $s = 1, 2$ est un indice de polarisation. La polarisation est orthogonale au vecteur d'onde $\vec{k}$ du photon, $\vec{e}_s\cdot\vec{k} = 0$, on doit intégrer sur tous les angles d'émission et sommer sur les polarisations. Il faut donc prendre la moyenne angulaire

$$\left\langle \sum_{s=1,2} \left(\hat{z}\cdot\vec{e}_s(\vec{k})\right)^2 \right\rangle = \frac{1}{4\pi}\int \mathrm{d}\Omega_k \sum_{s=1,2} \left(\hat{z}\cdot\vec{e}_s(\vec{k})\right)^2$$

Comme $\vec{e}_1$, $\vec{e}_2$ et $\hat{k}$ forment un trièdre

$$(\hat{z}\cdot\vec{e}_1)^2 + (\hat{z}\cdot\vec{e}_2)^2 = 1 - (\hat{z}\cdot\hat{k})^2 = \sin^2\theta$$

où θ est l'angle entre Oz et $\vec{k}$. La moyenne angulaire vaut

$$\langle\sin^2\theta\rangle = \frac{1}{4\pi}\int \mathrm{d}\Omega_k \sin^2\theta = \frac{2}{3}$$

Remaplaçant ω par ω_0 et tenant compte du facteur $1/2$ mentionné ci-dessus pour A, on déduit la probabilité d'émission spontanée pour une transition dipolaire électrique

$$\boxed{B' = \frac{2\hbar\omega_0^3}{3\pi^2c^3}\,\frac{A}{2} = \frac{4\omega_0^3}{3c^2}\left(\frac{d^2}{4\pi\varepsilon_0\hbar c}\right)} \tag{5.78}$$

Dans le cas de la physique atomique, un ordre de grandeur du moment dipolaire d est $d \sim q_e a$, où a est le rayon de l'orbite électronique, et on a l'estimation

$$B' \sim \alpha\frac{a^2\omega_0^3}{c^2} \sim \alpha^5\left(\frac{m_e c^2}{\hbar}\right) \tag{5.79}$$

où $\alpha = q_e^2/(4\pi\varepsilon_0\hbar c)$ est la constante de structure fine. Cette estimation est en accord avec (1.44), qui était fondé sur le calcul classique du rayonnement. Un calcul complet de B' sera donné au § 17.3.4, qui confirmera (5.78).

Bien que la RMN et les atomes à deux niveaux exhibent des analogies remarquables, et bien que le traitement mathématique soit identique, il n'en existe pas moins une importante différence de principe entre les deux cas. La mesure en RMN n'est pas une mesure projective du type défini en (4.7), mais elle utilise un signal *collectif* construit par un nombre macroscopique ($\sim 10^{20}$) de molécules. L'énergie du photon émis dans la transition entre les deux niveaux Zeeman du spin nucléaire, de l'ordre de $1\,\mu$eV, est bien trop faible pour être détectée pour une molécule isolée, et une autre conséquence est que l'émission spontanée est négligeable, en raison de sa dépendance en ω_0^3. Le détecteur RMN est une bobine enroulée autour de l'échantillon (figure 5.9). L'aimantation tournante induit une force électromotrice qui est détectée dans la bobine, et cette détection est un processus qui se décrit classiquement.

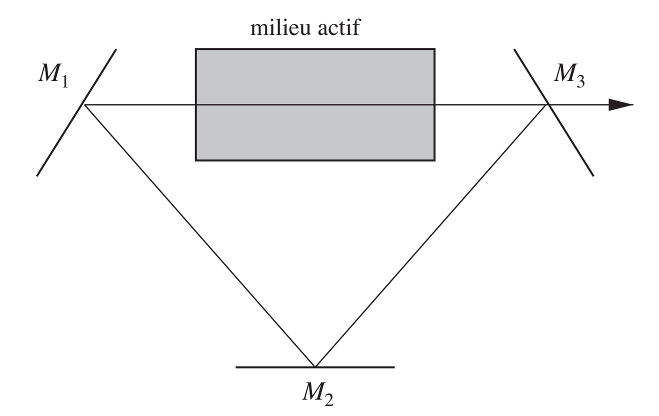


FIG. 5.18 – Schéma d'un laser en anneau. Les miroirs M_1 et M_2 sont parfaitement réfléchissants, M_3 transmet une fraction de la lumière incidente qui forme le faisceau laser.

5.4.2 Principes du laser

L'objectif de cette sous-section n'est évidemment pas de donner une vue générale des lasers et de leurs applications[22], mais de montrer comment les processus physiques décrits dans la sous-section précédente sont mis en œuvre dans cet instrument remarquable qu'est le laser. Il s'agit d'exploiter le processus d'émission stimulée, et, ainsi que nous l'avons vu, il est nécessaire d'obtenir une inversion de population. Un laser comprend en général :

- Un milieu actif qui joue le rôle d'amplificateur optique.
- Une source d'énergie qui permet d'exciter le milieu actif et de réaliser l'inversion de population par un processus dit de "pompage optique". Le pompage peut être électrique (cas des lasers à semi-conducteurs) ou optique (lampe ou autre laser).
- Une cavité résonante[23] linéaire ou en anneau (figure 5.18) qui sélectionne les modes amplifiés. Un des miroirs de la cavité est partiellement réfléchissant afin qu'une partie du rayonnement lumineux puisse être extraite de la cavité. Le laser est un milieu ouvert, et donc dissipatif. C'est en fait un convertisseur d'énergie de pompage en énergie électromagnétique.

Dans le cas de la cavité en anneau de la figure 5.18, le champ électrique doit être identique à lui-même lorsque la lumière a fait un tour complet, et les longueurs d'onde permises seront fixées par la stationnarité de la phase : la

22. La littérature sur les lasers est très vaste. Comme points de départ, on peut choisir le livre de Breiteneker et Treps [2010] ou l'article de Pocholle [2005]. Pour un exposé plus complet, mais focalisé sur les aspects de physique atomique fondamentale, voir Grynberg *et al.* [2010], chapitre 3.

23. La présence d'une cavité résonante n'est pas obligatoire. Il existe d'autres mécanismes pour sélectionner les modes amplifiés.

longueur optique L de la cavité doit être un multiple entier de la longueur d'onde λ

$$L = p\lambda \qquad p = 1,\ 2,\ \cdots \tag{5.80}$$

ou, en terme de fréquences,

$$\omega_p = \frac{2\pi c}{L}\, p \qquad p = 1,\ 2,\ \cdots \tag{5.81}$$

Il est généralement possible de sélectionner une longueur d'onde unique, ce qui donne un laser monomode. Grâce à l'utilisation de miroirs concaves, le faisceau laser possède une extension transverse finie et son ouverture angulaire à grande distance est, en raison de la diffraction, de l'ordre de λ/d, où d est la dimension transverse minimale, ou col, du faisceau.

Il existe plusieurs mécanismes permettant de réaliser l'inversion de population. Nous nous contenterons de décrire le mécanisme le plus simple, qui repose sur l'utilisation d'un système à quatre niveaux d'énergie : E_0 (fondamental), E_3 (excité), E_1 et E_2 (niveaux intermédiaires), voir la figure 5.19. Les atomes (ou molécules...) sont portés par pompage optique dans le niveau E_3 avec un taux w. Une relaxation rapide amène ensuite l'atome dans le niveau E_2, choisi tel que la relaxation de E_2 vers E_1 soit lente. Enfin, une relaxation rapide conduit de E_1 à E_0. Si on appelle Γ_i les taux de relaxation, on doit donc s'arranger pour avoir

$$\Gamma_1, \Gamma_3 \gg \Gamma_2 \tag{5.82}$$

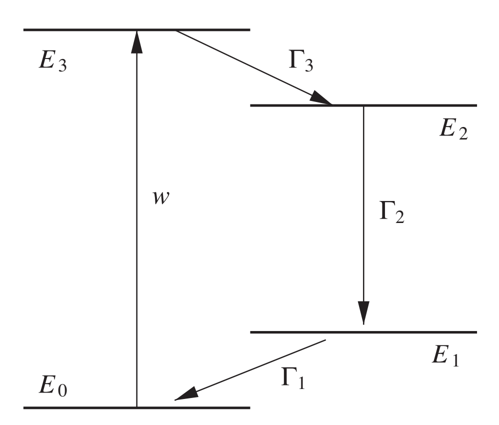


FIG. 5.19 – Schéma d'un mécanisme d'inversion de population à 4 niveaux.

En régime non saturé, on peut négliger l'émission stimulée ou l'absorption entre les niveaux E_2 et E_3, de sorte que les équations cinétiques pour les

populations N_i des niveaux E_i s'écrivent[24]

$$\begin{aligned}\frac{\mathrm{d}N_1}{\mathrm{d}t} &= \Gamma_2 N_2 - \Gamma_1 N_1 \\ \frac{\mathrm{d}N_2}{\mathrm{d}t} &= \Gamma_3 N_3 - \Gamma_2 N_2 \\ \frac{\mathrm{d}N_3}{\mathrm{d}t} &= wN_0 - \Gamma_3 N_3\end{aligned} \tag{5.83}$$

On vérifie la cohérence de ces équations en prenant leur somme, et compte tenu de $N_1 + N_2 + N_3 + N_4 = \text{cste}$, on obtient

$$\frac{\mathrm{d}N_0}{\mathrm{d}t} = -wN_0 + \Gamma_1 N_1 \tag{5.84}$$

qui est bien l'équation cinétique pour N_0. En régime permanent, $\mathrm{d}N_i/\mathrm{d}t = 0$ et on déduit de (5.83)

$$\frac{N_2}{N_1} = \frac{\Gamma_1}{\Gamma_2} \tag{5.85}$$

ce qui montre que la condition $\Gamma_1 \gg \Gamma_2$ entraîne l'inversion de population. En fait, l'émission stimulée est en compétition avec l'émission spontanée, et cette dernière est préjudiciable au fonctionnement du laser. Comme la probabilité d'émission spontanée est proportionnelle à ω_0^3 (voir 5.78), elle sera d'autant plus gênante que la longueur d'onde du laser sera courte. L'émission spontanée est environ huit fois plus importante dans le bleu que dans le rouge, et c'est pourquoi il est plus facile de construire un laser émettant dans le rouge que dans le vert ou le bleu. Malgré un rendement énergétique défavorable, un pointeur laser vert est construit en partant d'un laser émettant dans l'infrarouge, et le faisceau vert est obtenu grâce à une conversion optique depuis l'infrarouge en vert par doublement de fréquence dans un milieu non-linéaire.

La condition de stationnarité de l'onde laser dans la cavité est à l'origine des propriétés de cohérence temporelle (ou longitudinale) et de cohérence spatiale (ou transverse) de la lumière laser, qui sont à la base des propriétés remarquables de cette lumière. Pour définir la cohérence temporelle, remarquons que si l'onde lumineuse était une sinusoïde parfaite de fréquence ω, on aurait en un point d'espace fixé un comportement en $\cos(\omega t - \phi)$. En fait, ϕ est une fonction aléatoire du temps, $\phi(t)$. Si $\phi(t)$ a une certaine valeur à $t = 0$, la mémoire de cette valeur sera perdue au bout d'un temps τ, le temps de corrélation de la phase, ou temps de cohérence, et $c\tau$ est la longueur de cohérence longitudinale. On peut, en principe, atteindre un temps de cohérence limité uniquement par l'émission spontanée, et, en pratique, des temps de cohérence de l'ordre de la dizaine de μs (et donc des longueurs de cohérence de plusieurs kilomètres) sont faciles à atteindre. Un argument analogue est utilisé pour

24. Les équations cinétiques complètes sont données, par exemple, par Pocholle [2005].

définir la cohérence entre deux points dans le plan perpendiculaire au faisceau, ce qui donne la cohérence transverse. La cohérence transverse permet à la largeur du faisceau d'approcher les limites de la diffraction, c'est-à-dire de focaliser le faisceau sur une tache dont les dimensions sont de l'ordre de la longueur d'onde. C'est aussi cette propriété qui permet d'obtenir des faisceaux lumineux dont la divergence approche la seconde d'arc en les dilatant jusqu'à un diamètre de 15 cm, et on peut, par exemple, former sur la Lune une tache de quelques km de diamètre. Grâce à la réflexion de ce faisceau par un miroir, il est possible de mesurer la distance Terre-Lune avec une précision de 1 mm.

Si l'on compare un laser à une source de lumière conventionnelle, par exemple un laser YAG doublé en fréquence émettant dans le vert une puissance de 15 W, et une ampoule à incandescence de 150 W (exemple emprunté à Grynberg *et al.* [2010]), le laser coûte environ 40 000 euros et la lampe à incandescence 1 euro : il semble que le laser représente une technologie sophistiquée et onéreuse pour une puissance relativement faible ! Mais l'avantage du laser devient décisif quand on essaie de concentrer la puissance sur une petite surface ou un intervalle de temps très court. Compte tenu des lois de l'optique, il est impossible d'obtenir un flux lumineux (une puissance par unité de surface) de plus de 500 $\mathrm{W/cm^2}$ avec l'ampoule. Au contraire, en raison de la cohérence spatiale de la lumière laser, la puissance du laser YAG peut être focalisée sur une surface dont la taille est limitée uniquement par la diffraction, c'est-à-dire par la longueur d'onde émise, $\lambda \sim 0.5\,\mu$m. Le flux lumineux peut alors atteindre $10^9\,\mathrm{W/cm^2}$, soit un gain de plus d'un million par rapport à l'ampoule. La cohérence temporelle permet de concentrer la puissance sur des intervalles de temps très courts et avec un taux de répétition élevé grâce à la technique dite de synchronisation de modes. Par exemple, un laser titane-saphir à modes synchronisés permet d'atteindre une puissance de $10^{14}\,$W sur des intervalles de temps de l'ordre de la femtoseconde ($10^{-15}\,$s, lasers femtosecondes).

Une façon de comprendre la différence entre la lumière laser et la lumière d'une source classique est de comparer le nombre de photons dans un mode. Dans le cas de la lumière du Soleil (source lumineuse correspondant à une température de 5 500 °C), le nombre moyen de photons par mode dans le visible est de l'ordre de 1/100. Au contraire, le nombre moyen de photons par mode dans un laser peut atteindre des valeurs de 10^9. C'est ce très grand nombre de photons par mode qui permet de concentrer l'énergie sur des petites surfaces ou des intervalles de temps très courts.

Il existe une grande variété de lasers : lasers à néodyme (YAG), lasers saphir-titane, lasers de puissance à CO_2, ... Les lasers les plus répandus dans la vie courante sont les lasers à semi-conducteurs, ou diodes laser. Dans ce cas, le processus de base est la recombinaison d'une paire électron-trou, et les dimensions de la cavité résonante sont de l'ordre de quelques centaines de μm. Le laser a été inventé il y a tout juste 50 ans, mais on ne pourrait plus imaginer aujourd'hui un monde sans lasers ! Le laser est bien sûr omniprésent dans

les laboratoires de recherche, mais aussi dans le monde industriel : découpe industrielle, gravure, soudure, et le monde médical : micro-chirurgie. Les lasers sont utilisés pour le positionnement : télémétrie, gyrolasers, etc. Les diodes laser sont fondamentales pour la lecture des CD, des DVD, des codes barres et pour les télécommunications par fibre optique. Les lasers femtosecondes ont une puissance de crête pouvant atteindre 100 TW : le laser mégajoule à Bordeaux pourrait déclencher la fusion thermonucléaire inertielle contrôlée et représente une alternative à ITER. Bref, le laser est devenu un outil universel.

5.4.3 Franges de Ramsey et principe des horloges atomiques

La dernière application des sytèmes à deux niveaux étudiée dans ce chapitre est celle des horloges atomiques. Une horloge repose toujours sur un phénomène périodique, que ce soit le mouvement du balancier de l'horloge de nos grand-mères ou la vibration d'un cristal de quartz des montres actuelles. La seconde est aujourd'hui *définie* par la fréquence ν_0 d'une transition entre deux niveaux hyperfins (§ 15.2.4) de l'isotope 133 de l'atome de césium, qui sert donc d'étalon universel de temps. Si E_g et E_e ($E_e > E_g$) sont les énergies de ces deux niveaux, alors la fréquence ν_0 est donnée par

$$h\nu_0 = E_e - E_g$$

La fréquence ν_0 de la vibration électromagnétique émise dans cette transition est *par définition* de 9 192 631 770 Hz, environ 10^{10} Hz. Pour fabriquer une horloge, il faut être capable de comparer la fréquence du signal de l'horloge à celle de la transition atomique. Il est nécessaire d'"interroger" les atomes, c'est-à-dire comparer le signal de la transition au signal de l'horloge. Ceci est illustré sur la figure 5.20, où sont comparés le signal de l'horloge et celui de la transition. Si l'on observe les deux signaux suffisamment longtemps, on met en évidence la différence de fréquence, alors qu'une comparaison sur un temps plus court ne permet pas de déceler de différence : la précision de l'horloge sera d'autant meilleure que la durée T de comparaison sera plus longue. Cette comparaison entre le signal de l'horloge et le signal de référence permet de corriger les erreurs, et ainsi d'asservir le signal de l'horloge au signal de référence. Il faut bien sûr éliminer du mieux possible tous les effets parasites qui pourraient influencer la vibration émise. La méthode utilisée pour asservir le signal de l'horloge au signal atomique est celle des *franges de Ramsey*.

Nous allons modéliser la transition dans l'atome de césium par un atome à deux niveaux $|g\rangle$ (fondamental) et $|e\rangle$ (excité), avec une fréquence de résonance $\omega_0 = 2\pi\nu_0$, $\hbar\omega_0 = E_e - E_g$. L'atome initialement dans l'état $|g\rangle$ est envoyé au temps $t = 0$ avec une vitesse v dans une cavité C_1, où il est soumis à un champ électrique périodique de fréquence ω (le signal de l'horloge). Après un temps $T = L/v$, où L est la distance entre les deux cavités, il traverse une seconde cavité C_2 *identique* à C_1. Les cavités C_1 et C_2 sont reliées à une

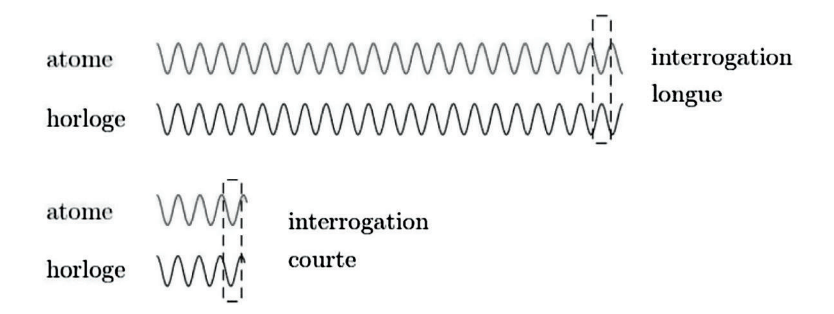


FIG. 5.20 – Comparaison entre le signal de l'atome (sinusoïde du dessus) et celui de l'horloge (sinusoïde du dessous). Les deux vibrations coïncident au temps $t = 0$. Si l'on observe les deux signaux suffisamment longtemps, on met en évidence une différence (sinusoïdes supérieures), alors que les signaux sont pratiquement confondus sur un intervalle de temps court (sinusoïdes inférieures).

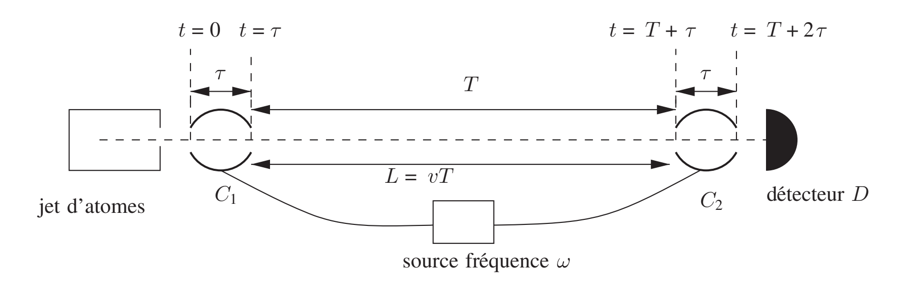


FIG. 5.21 – Principe d'une horloge atomique. Le détecteur D est déclenché uniquement par les atomes dans l'état $|g\rangle$.

même source S et leurs champs électriques sont en phase (figure 5.21) ; les atomes passent un temps τ dans chacune des cavités, avec $\tau \ll T$. Ils sont détectés immédiatement à la sortie de C_2 par un détecteur D, qui se déclenche *uniquement si les atomes sont dans l'état* $|g\rangle$.

Les états $|g\rangle$ et $|e\rangle$ sont choisis comme états propres de σ_z

$$\sigma_z|g\rangle = |g\rangle \qquad \sigma_z|e\rangle = -|e\rangle$$

Le hamiltonien de l'atome s'écrit au passage par les cavités, c'est-à-dire pour $0 \leq t \leq \tau$ et pour $T + \tau \leq t \leq T + 2\tau$

$$H(t) = -\frac{1}{2}\,\hbar\omega_0\sigma_z - \hbar\omega_1\sigma_x\,\cos\omega t \tag{5.86}$$

tandis que la fréquence de Rabi $\omega_1 = 0$ pour $T \leq t \leq T + \tau$. D'après (5.34), le hamiltonien $\hat{H}$ dans le référentiel tournant est indépendant du temps, et

pour $0 \leq t \leq \tau$ ainsi que pour $T+\tau \leq t \leq T+2\tau$

$$\hat{H} = \frac{1}{2}\,\hbar\delta\sigma_z - \frac{1}{2}\,\hbar\omega_1\sigma_x \qquad \hat{U}(t) = \mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar} \tag{5.87}$$

où $\delta = \omega - \omega_0$ est le désaccord entre la fréquence ω de la source et la fréquence de référence ω_0. La fréquence de Rabi ω_1 est proportionnelle au champ électrique qui règne dans les cavités C_1 et C_2, et $\omega_1 = 0$ en dehors de ces cavités. D'après le § 5.2.1, la loi d'évolution du vecteur d'état $|\hat{\varphi}(t)\rangle$ de l'atome est donnée dans le référentiel tournant par

$$|\hat{\varphi}(t)\rangle = \mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar}|\hat{\varphi}(t=0)\rangle = \hat{U}(t)|\hat{\varphi}(t=0)\rangle$$

On se place d'abord à la résonance : $\delta = 0$, et on s'arrange pour que l'atome initialement dans l'état $|g\rangle$ traverse C_1 en un temps τ tel que $\omega_1\tau = \pi/2$; la condition $\omega_1\tau/2 = \pi/4$ correspond à une impulsion $\pi/2$ qui transforme $|g\rangle$ en $(|g\rangle + \mathrm{i}|e\rangle)/\sqrt{2}$.

Pendant l'intervalle de temps T où l'atome se propage entre C_1 et C_2, $\hat{U}(t) = I$: *dans le référentiel tournant*, l'atome n'évolue pas entre C_1 et C_2. L'atome met un temps τ à traverser C_2, soit une nouvelle impulsion $\pi/2$

$$|g\rangle \to \frac{1}{\sqrt{2}}(|g\rangle + \mathrm{i}|e\rangle) \qquad |e\rangle \to \frac{1}{\sqrt{2}}(\mathrm{i}|g\rangle + |e\rangle) \tag{5.88}$$

ce qui implique $(|g\rangle + \mathrm{i}|e\rangle)/\sqrt{2} \to \mathrm{i}|e\rangle$. L'atome se retrouve dans l'état excité car deux impulsions $\pi/2$ successives sont équivalentes à une impulsion π, soit une rotation de π sur la sphère de Poincaré-Bloch. Dans ces conditions, les atomes arrivent dans l'état $|e\rangle$ et le détecteur n'est pas déclenché. Cette méthode, dite des *franges de Ramsey*, repose sur le même principe que l'interféromètre de Mach-Zehnder. La cavité C_1 joue le rôle de la première lame séparatrice et C_2 celle de la seconde. Dans l'interféromètre de Mach-Zehnder, les trajets des photons correspondent à des propagations spatiales différentes, et dans la méthode des franges de Ramsey, les trajets des atomes sont situés dans l'espace de Hilbert des deux états atomiques.

Pour une étude complète, il faut maintenant se placer en dehors de la résonance, $\delta \neq 0$. À la sortie de C_1, l'état de l'atome est $\hat{U}(\tau)|g\rangle$, avec (voir (5.42))

$$\hat{U}(\tau) = I\cos\frac{\Omega\tau}{2} + \mathrm{i}\left(\frac{\omega_1}{\Omega}\,\sigma_x - \frac{\delta}{\Omega}\,\sigma_z\right)\sin\frac{\Omega\tau}{2} \qquad \Omega = \sqrt{\omega_1^2 + \delta^2} \tag{5.89}$$

Pour $\delta \ll \omega_1$,

$$\Omega = \sqrt{\omega_1^2 + \delta^2} \simeq \omega_1\left(1 + \frac{\delta^2}{2\omega_1^2}\right) \qquad \frac{\omega_1}{\Omega} = 1 - \frac{\delta^2}{2\omega_1^2}$$

ce qui entraîne

$$\cos\frac{\Omega\tau}{2} = \frac{1}{\sqrt{2}}\left(1 + O\left(\frac{\delta}{\omega_1}\right)^2\right) \qquad \sin\frac{\Omega\tau}{2} = \frac{1}{\sqrt{2}}\left(1 + O\left(\frac{\delta}{\omega_1}\right)^2\right)$$

Si l'on néglige les termes de l'ordre de $(\delta/\omega_1)^2$, l'état $|\chi\rangle$ de l'atome à la sortie de C_1 est

$$|\chi\rangle = \frac{1}{\sqrt{2}}\left[\left(1 - \frac{\mathrm{i}\delta}{\omega_1}\right)|g\rangle + \mathrm{i}|e\rangle\right]$$

De plus, la propagation de l'atome entre C_1 et C_2 est modifiée, car pour $\tau \leq t \leq T + \tau$ nous avons dans le référentiel tournant

$$\hat{H} = \frac{1}{2}\hbar\delta\sigma_z \qquad \hat{U}(t) = \mathrm{e}^{-\mathrm{i}\delta\sigma_z t/2}$$

Négligeant les terme en (δ/ω_1) dans l'expression de $|\chi\rangle$, à l'entrée de C_2 l'état de l'atome est

$$\mathrm{e}^{-\mathrm{i}\delta\sigma_z T/2}\frac{1}{\sqrt{2}}(|g\rangle + \mathrm{i}|e\rangle) = \frac{1}{\sqrt{2}}\left(\mathrm{e}^{-\mathrm{i}\delta T/2}|g\rangle + \mathrm{i}\mathrm{e}^{\mathrm{i}\delta T/2}|e\rangle\right)$$

L'action de C_2 est identique à celle du cas résonant et d'après (5.88) l'atome à la sortie de C_2 se trouve dans l'état

$$\frac{1}{2}\left(\mathrm{e}^{-\mathrm{i}\delta T/2}(|g\rangle + \mathrm{i}|e\rangle) + \mathrm{i}\mathrm{e}^{\mathrm{i}\delta T/2}(|e\rangle + \mathrm{i}|g\rangle)\right) = \mathrm{i}\left(-\sin\frac{\delta T}{2}\,|g\rangle + \cos\frac{\delta T}{2}\,|e\rangle\right) \tag{5.90}$$

La probabilité de déclencher D est donc

$$\mathsf{p}(T) = \sin^2\frac{\delta T}{2} \tag{5.91}$$

L'asservissement de la fréquence ω à ω_0 consiste à rendre la probabilité de détection (5.91) aussi petite que possible. Si le détecteur enregistre un maximum d'atomes, cela veut dire que $\delta T = \pi$: la méthode détecte donc un désaccord $\delta \sim 1/T$. Plus le temps de parcours est long, et meilleure sera la précision de l'horloge. Cette précision atteint 10^{-13} en valeur relative, ce qui correspond à une dérive d'une seconde tous les 300 000 ans.

5.5 Exercices

5.5.1 Base orthonormée de vecteurs propres

Vérifier par un calcul explicite que les vecteurs $|\chi_s\rangle$ (5.12) forment une base orthonormée : $\langle\chi_{s'}|\chi_s\rangle = \delta_{s's}$.

5.5.2 Moment dipolaire électrique du formaldéhyde

1. On se propose de modéliser le comportement des deux électrons π de la double liaison de la molécule de formaldéhyde H_2-C=O. L'oxygène étant plus électronégatif que le carbone, montrer que le hamiltonien d'un électron prend la forme

$$\begin{pmatrix} E_{\mathrm{C}} & -A \\ -A & E_{\mathrm{O}} \end{pmatrix}$$

avec $E_O < E_C$, où $E_C(E_O)$ représente l'énergie d'un électron localisé sur l'atome de carbone (d'oxygène).

2. On définit

$$B = \frac{1}{2}(E_C - E_O) > 0$$

et l'angle θ par

$$B = \sqrt{A^2 + B^2}\cos\theta \qquad A = \sqrt{A^2 + B^2}\sin\theta$$

Calculer en fonction de θ la probabilité de trouver un électron π localisé sur l'atome de carbone ou d'oxygène.

3. On *admet* que le moment dipolaire électrique d du formaldéhyde est dû uniquement à la distribution de charge sur l'axe C=O. Exprimer ce moment dipolaire en fonction de la distance l entre les atomes de carbone et d'oxygène, la charge du proton q_p et θ. Les valeurs expérimentales sont $l = 0.121$ nm et $d = q_p \times 0.040$ nm.

5.5.3 Le butadiène

Le butadiène C_4H_6 possède une structure linéaire (figure 5.22). Son ossature $(C_4H_6)^{4+}$ formée avec des électrons σ comporte quatre atomes de carbone numérotés de $n = 1$ à $n = 4$. L'état d'un électron π localisé au voisinage de l'atome de carbone n°n est désigné par $|\varphi_n\rangle$. Il est commode de généraliser à une chaîne linéaire comportant un nombre N d'atomes de carbone, et donc de numéroter les atomes $n = 1, \ldots, N$. Le hamiltonien d'un électron π agit sur l'état $|\varphi_n\rangle$ de la façon suivante

$$\begin{aligned} H|\varphi_n\rangle &= E_0|\varphi_n\rangle - A(|\varphi_{n-1}\rangle + |\varphi_{n+1}\rangle) \quad \text{si} \quad n \neq 1,\ N \\ H|\varphi_1\rangle &= E_0|\varphi_1\rangle - A|\varphi_2\rangle \\ H|\varphi_N\rangle &= E_0|\varphi_N\rangle - A|\varphi_{N-1}\rangle \end{aligned}$$

A est une constante positive. On remarque que les états $|\varphi_1\rangle$ et $|\varphi_N\rangle$ jouent un rôle particulier, car contrairement au cas du benzène, il n'y a pas de symétrie cyclique.

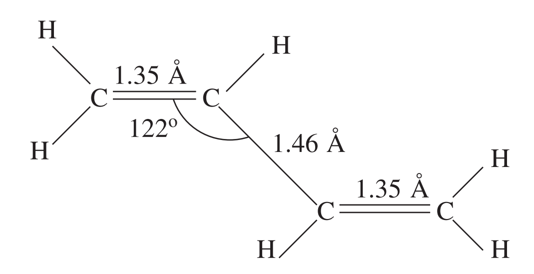


Fig. 5.22 – Formule chimique du butadiène.

1. Écrire la forme matricielle explicite de H dans la base $|\varphi_n\rangle$ pour $N = 4$.

2. L'état le plus général pour un électron π est

$$|\chi\rangle = \sum_{n=1}^{N} c_n |\varphi_n\rangle$$

Pour adapter la méthode utilisée dans le cas de la symétrie cyclique, on introduit deux états fictifs $|\varphi_0\rangle$ et $|\varphi_{N+1}\rangle$ et deux composantes $c_0 = c_{N+1} = 0$, ce qui permet de récrire $|\chi\rangle$

$$|\chi\rangle = \sum_{n=0}^{N+1} c_n |\varphi_n\rangle$$

Montrer que l'action de H sur l'état $|\chi\rangle$ s'écrit

$$H|\chi\rangle = E_0|\chi\rangle - A\sum_{n=1}^{N}(c_{n-1} + c_{n+1})|\varphi_n\rangle$$

3. En s'inspirant de la méthode utilisée dans le cas de la symétrie cyclique on cherche c_n sous la forme

$$c_n = \frac{c}{2\mathrm{i}}\left(\mathrm{e}^{\mathrm{i}n\delta} - \mathrm{e}^{-\mathrm{i}n\delta}\right)$$

ce qui assure que $c_0 = 0$. Montrer que l'on doit choisir

$$\delta = \frac{\pi s}{N+1} \qquad s = 1, \ldots, N$$

si l'on veut également avoir $c_{N+1} = 0$.

4. Montrer que les valeurs propres de H sont étiquetées par l'entier s

$$E_s = E_0 - 2A\cos\frac{\pi s}{N+1}$$

et donner l'expression des vecteurs propres $|\chi_s\rangle$ correspondants. Montrer que la constante de normalisation c vaut $\sqrt{2/(N+1)}$ (suggestion : *cf.* (5.15)).

5. Dans le cas du butadiène $N = 4$, donner les valeurs numériques de E_s et des composantes des vecteurs propres. Montrer que l'énergie de l'état fondamental de l'ensemble des 4 électrons π est

$$E_0 \simeq 4(E_0 - A) - 0.48A$$

Le gain dû à la délocalisation des électrons π sur l'ensemble de la chaîne est-il important par rapport à la formule chimique de la figure 5.22 ? Dessiner qualitativement la probabilité de présence des électrons pour $s = 1$ et $s = 2$.

6. Quelle serait l'énergie de l'état fondamental d'une hypothétique molécule cyclique (c'est-à-dire ayant la forme d'un carré) C_4H_4 ?

7. On définit l'ordre d'une liaison l entre deux atomes de carbone n et $n+1$ par

$$l = 1 + \sum_s \langle\varphi_n|\chi_s\rangle\langle\chi_s|\varphi_{n+1}\rangle$$

où la somme porte sur les états $|\chi_s\rangle$ occupés par les électrons π. Le facteur 1 correspond aux électrons σ. Montrer que l'ordre de la liaison est bien $l = 2$ pour l'éthylène. Calculer l'ordre des liaisons pour le benzène et pour les différentes liaisons du butadiène et commenter les résultats. Pourquoi la liaison centrale du butadiène est-elle plus courte qu'une liaison simple (1.46 Å au lieu de 1.54 Å) ?

5.5.4 Vecteurs propres du hamiltonien (5.22)

Montrer que dans le cas où le champ électrique est indépendant du temps, et lorsque $d\mathcal{E}/A \ll 1$, le vecteur propre normalisé de H correspondant à la valeur propre $E_0 - A$ est donné à l'ordre $d\mathcal{E}/A$ par

$$|\chi'_+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 - d\mathcal{E}/(2A) \\ 1 + d\mathcal{E}/(2A) \end{pmatrix}$$

Quel est l'autre vecteur propre ?

5.5.5 L'ion moléculaire H_2^+

L'ion moléculaire H_2^+ est formé de deux protons et d'un électron. Les deux protons sont situés sur un axe pris comme axe des x, à des abscisses respectives $-r/2$ et $r/2$. Ils seront supposés fixes (approximation de Born-Oppenheimer, voir le § 16.3.1).

1. En supposant l'électron sur l'axe des x, exprimer son énergie potentielle $V(x)$ en fonction de sa position x et de $e^2 = q_e^2/(4\pi\varepsilon_0)$ où q_e est la charge de l'électron, et la tracer qualitativement.

2. Si les deux protons sont très éloignés, $r \gg l$, l'électron est soit localisé au voisinage du proton de droite (état $|\varphi_1\rangle$), soit au voisinage du proton de gauche (état $|\varphi_2\rangle$). On suppose que ces états correspondent tous deux à l'état fondamental de l'atome d'hydrogène d'énergie

$$E_0 = -\frac{1}{2}\,\frac{m_e e^4}{\hbar^2} = -\frac{e^2}{2a_0}$$

où m_e est la masse de l'électron et a_0 le rayon de Bohr : $a_0 = \hbar^2/(m_e e^2)$. Quelle est l'échelle de longueur pertinente l dans la relation $r \gg l$?

3. On considère l'ion H_2^+ comme un système à deux niveaux d'états de base $\{|\varphi_1\rangle, |\varphi_2\rangle\}$, $\langle\varphi_i|\varphi_j\rangle = \delta_{ij}$. Justifier la forme du hamiltonien où l'on choisira $A > 0$

$$H = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix}$$

Quels sont les états propres $|\chi_+\rangle$ et $|\chi_-\rangle$ de H et les énergies E_+ et E_-, $E_+ < E_-$, correspondantes ? Dessiner qualitativement les fonctions d'onde $\chi_\pm(x) = \langle x|\chi_\pm\rangle$ de l'électron sur l'axe des x.

4. Le paramètre A est une fonction de la distance r entre les protons : $A(r)$. Justifier le fait que A est une fonction décroissante de r et que $\lim_{r\to\infty} A(r) = 0$. L'énergie de l'électron est donc une fonction de r, $E_\pm(r)$.

5. Montrer que pour trouver l'énergie totale de l'ion $E'_\pm(r)$, on doit ajouter un terme $+e^2/r$. Quelle est l'origine physique de ce terme ?

6. On paramétrise $A(r)$ par

$$A(r) = c\, e^2 \exp\left(-\frac{r}{b}\right)$$

où b est une longueur et c l'inverse d'une longueur. Donner l'expression des deux niveaux d'énergie E'_+ et E'_- de l'ion. Soit

$$\Delta E(r) = E'_+(r) - E_0$$

la différence d'énergie entre l'état fondamental de l'ion et celui de l'atome d'hydrogène. Montrer que $\Delta E(r)$ peut passer par un minimum pour une valeur $r = r_0$ et en déduire

$$\Delta E(r_0) = \frac{e^2}{r_0}\left(1 - \frac{b}{r_0}\right)$$

Quelle condition doit-on avoir sur b et r_0 pour que l'ion H_2^+ soit un état lié ?

7. Les valeurs expérimentales sont $r_0 \simeq 2a_0$ et $\Delta E(r_0) \simeq E_0/5 = -e^2/(10a_0)$. En déduire b et c en fonction de a_0.

5.5.6 Compléments sur la RMN

1. Obtenir la formule (5.43) en examinant la figure 5.7b décrivant la rotation du spin dans le référentiel tournant. Suggestion : utiliser l'équation (7.23) en décomposant $\vec{b}$ suivant un vecteur $\hat{n} \parallel \vec{B}_{\text{eff}}$ et un vecteur perpendiculaire, avec $\theta = \Omega t$.

2. On suppose que le champ de radio-fréquences $\vec{B}_1$ est de la forme

$$\vec{B}_1(t) = 2\hat{x}B_1 \cos\omega t$$

Montrer que le hamiltonien dans le référentiel tournant est

$$\hat{H}(t) = \frac{1}{2}\,\hbar\delta\sigma_z - \frac{1}{2}\,\hbar\omega_1\sigma_x + \hbar\omega_1\left[\sigma_+ \mathrm{e}^{-2\mathrm{i}\omega t} + \sigma_- \mathrm{e}^{2\mathrm{i}\omega t}\right]$$

Que peut-on dire des termes entre crochets à l'approximation séculaire ? Interprétation physique : on récrit $\vec{B}_1$

$$\vec{B}_1(t) = B_1(\hat{x}\cos\omega t - \hat{y}\sin\omega t) + B_1(\hat{x}\cos\omega t + \hat{y}\sin\omega t)$$

Quel est le terme éliminé par l'approximation séculaire ? Pourquoi ?

5.6 Bibliographie

Pour la chimie quantique élémentaire, on pourra consulter Feynman *et al.* [1965], volume III, chapitre 15, Goodrich [1972], chapitre 2, Gatz [1971], chapitres 10 à 12. Les systèmes à deux niveaux avec interactions résonantes et quasi-résonantes sont traités par Feynman *et al.* [1965], volume III, chapitres 8 et 9 ou par Cohen-Tannoudji *et al.* [1973], chapitre IV. On trouvera une excellente introduction à la RMN par exemple dans Akitt [1992] ou dans Levitt [2001]. L'interaction d'un atome à deux niveaux avec un champ électromagnétique est traitée à un niveau avancé par Grynberg *et al.* [2010], chapitre 2. Le lecteur trouvera des détails complémentaires sur l'ion moléculaire H_2^+ dans Cohen-Tannoudji *et al.* [1973], complément G_{XI}.

Chapitre 6

Mathématiques de la mécanique quantique II : dimension infinie

Nous avons vu au chapitre 4 que les relations de commutation canoniques imposaient d'utiliser un espace des états de dimension infinie, dont le traitement rigoureux exigerait un outillage mathématique important. Heureusement, les physiciens peuvent en général se contenter de transposer au cas de la dimension infinie les résultats démontrés dans le cas de la dimension finie, avec des modifications simples que nous allons indiquer, sans avoir à se lancer dans des mathématiques trop complexes. Néanmoins, il n'est pas inutile d'être conscient des impasses sur la rigueur dont les physiciens sont coutumiers, afin d'éviter d'éventuelles mauvaises surprises.

L'objectif de ce chapitre est donc, d'une part, d'illustrer sur quelques exemples concrets les nouveautés apportées par la dimension infinie et, d'autre part, de donner des règles de calcul pratiques et en particulier d'écrire la décomposition spectrale des opérateurs hermitiens et unitaires. Les explications mathématiques sont un peu plus détaillées que celles données habituellement dans les manuels de mécanique quantique. Le lecteur intéressé uniquement par les aspects pratiques peut passer directement à la section 6.3, où sont rassemblés les résultats essentiels pour la suite.

6.1 Espaces de Hilbert

6.1.1 Définitions

L'espace des états de la mécanique quantique est un espace de Hilbert $\mathcal{H}$, en général de dimension infinie. La définition axiomatique d'un espace de Hilbert est la suivante.

1. C'est un espace vectoriel qui, pour les besoins de la mécanique quantique, est défini sur le corps des complexes. Les vecteurs de cet espace sont notés $|\varphi\rangle$.

2. Cet espace est muni d'un produit scalaire défini positif ; si $|\varphi\rangle$ et $|\chi\rangle$ sont deux vecteurs, ce produit scalaire est noté $\langle\chi|\varphi\rangle$ et il vérifie

$$\langle\chi|\varphi\rangle = (\langle\varphi|\chi\rangle)^* \tag{6.1}$$
$$\langle\chi|\varphi + \lambda\psi\rangle = \langle\chi|\varphi\rangle + \lambda\langle\chi|\psi\rangle \tag{6.2}$$
$$\langle\varphi|\varphi\rangle = ||\varphi||^2 = 0 \Longleftrightarrow |\varphi\rangle = 0 \tag{6.3}$$

λ étant un nombre complexe arbitraire ; $||\varphi||$ désigne la norme de $|\varphi\rangle$.

3. $\mathcal{H}$ est un *espace complet*, c'est-à-dire un espace où toute suite de Cauchy a une limite : si une suite de vecteurs $|\varphi^{(l)}\rangle$ de $\mathcal{H}$ est telle que $||\varphi^{(l)} - \varphi^{(m)}|| \to 0$ pour $l, m \to \infty$, alors il existe un vecteur $|\varphi\rangle$ de $\mathcal{H}$ tel que $||\varphi^{(l)} - \varphi|| \to 0$ pour $l \to \infty$. Si $\mathcal{H}$ n'est pas complet, on peut toujours lui rajouter les vecteurs limites de suites de Cauchy et le rendre complet[1].

4. Un espace de Hilbert est *caractérisé par sa dimension* : tous les espaces de même dimension sont isomorphes. La dimension d'un espace de Hilbert peut être finie et égale à N, elle peut être infinie dénombrable, ou bien non dénombrable. Les espaces de Hilbert qui interviennent en mécanique quantique sont soit de dimension finie, soit de dimension infinie dénombrable.

Nous avons étudié en détail au chapitre 2 les espaces de Hilbert de dimension finie. Si la dimension est N, il faut N vecteurs unitaires orthogonaux $|n\rangle$, $n = 1, \ldots, N$ pour former une base orthonormée : $|1\rangle, |2\rangle, \ldots, |n\rangle, \ldots, |N\rangle$. Dans le cas dénombrable, il existe une suite dénombrable de vecteurs unitaires orthogonaux $|1\rangle, |2\rangle, \ldots, |n\rangle, \ldots$ formant une base de $\mathcal{H}$: tout vecteur de $\mathcal{H}$ peut s'écrire comme combinaison linéaire de ces vecteurs de base

$$|\varphi\rangle = \sum_{n=1}^{\infty} c_n|n\rangle \tag{6.4}$$

mais contrairement au cas de la dimension finie, toute combinaison de la forme (6.4) n'est pas un vecteur de $\mathcal{H}$! En effet, le carré de la norme de $|\varphi\rangle$ est donné par

$$||\varphi||^2 = \sum_{n=1}^{\infty} |c_n|^2 \tag{6.5}$$

et (6.4) ne définit un vecteur que si cette norme est finie : la série dans (6.5) doit être une série convergente

$$\sum_{n=1}^{\infty} |c_n|^2 < \infty$$

1. Cet axiome est donc, en fait, un peu superflu. Il est automatiquement vérifié dans le cas de la dimension finie.

Dans ces conditions, quel que soit $\varepsilon > 0$, il existe un entier N tel que le vecteur $|\varphi_N\rangle$ défini par la combinaison suivante *finie* de vecteurs de base

$$|\varphi_N\rangle = \sum_{n=1}^{N} c_n |n\rangle$$

vérifie

$$||\varphi - \varphi_N||^2 = \sum_{n=N+1}^{\infty} |c_n|^2 \leq \varepsilon \tag{6.6}$$

Autrement dit, il est possible d'approcher $|\varphi\rangle$ par un vecteur $|\varphi_N\rangle$ dont la norme diffère arbitrairement peu de celle de $|\varphi\rangle$. On peut maintenant approcher les c_n par des nombres rationnels, et on voit qu'il est possible de construire dans $\mathcal{H}$ une suite dénombrable de vecteurs qui soit dense[2] dans $\mathcal{H}$. Cette propriété, commune aux espaces de dimension finie et dénombrable, s'appelle la *séparabilité* de l'espace de Hilbert : les espaces de Hilbert de la mécanique quantique sont séparables.

La convergence définie par (6.6) est la *convergence en norme*, aussi appelée *convergence forte* : on dit qu'une suite de vecteurs $|\varphi^{(l)}\rangle$ converge en norme vers $|\varphi\rangle$ pour $l \to \infty$ si quel que soit $\varepsilon > 0$, il existe un entier N tel que pour $l \geq N$

$$||\varphi - \varphi^{(l)}|| \leq \varepsilon \qquad \forall\, l \geq N \tag{6.7}$$

Il existe un autre type de convergence, la *convergence faible* : une suite de vecteurs $|\varphi^{(l)}\rangle$ converge faiblement vers $|\varphi\rangle$ si pour tout vecteur $|\chi\rangle$ de $\mathcal{H}$

$$\lim_{l\to\infty} \langle\varphi^{(l)}|\chi\rangle = \langle\varphi|\chi\rangle \tag{6.8}$$

Nous n'aurons pas à nous servir de la convergence faible[3], mais l'existence de cette convergence permet d'illustrer une différence avec la dimension finie : les deux convergences sont identiques pour un espace de dimension finie, mais non pour un espace de dimension infinie. La convergence forte implique la convergence faible, mais non l'inverse (exercice 6.4.1).

6.1.2 Réalisations d'espaces séparables et de dimension infinie

Tous les espaces de Hilbert séparables et de dimension infinie sont isomorphes ; cependant, les réalisations concrètes peuvent *a priori* sembler différentes et il est intéressant de pouvoir les identifier. Nous allons définir successivement les espaces $\ell^{(2)}$, $L^{(2)}[a,b]$ et $L^{(2)}(\mathbb{R})$, qui sont tous séparables et de dimension infinie.

2. Un ensemble de vecteurs $\{|\varphi^{(\alpha)}\rangle\}$ est dense dans $\mathcal{H}$, si pour tout $\varepsilon > 0$ et pour tout vecteur $|\varphi\rangle$ de $\mathcal{H}$, on peut trouver un $|\varphi^{(\alpha)}\rangle$ tel que $||\varphi - \varphi^{(\alpha)}|| < \varepsilon$.

3. Elle intervient par exemple dans certains problèmes de théorie quantique des champs.

(i) *Espace* $\ell^{(2)}$. Un vecteur $|\varphi\rangle$ est défini par une suite infinie de nombres complexes $c_1, \ \ldots, c_n \ldots$ telle que

$$||\varphi||^2 = \sum_{n=1}^{\infty} |c_n|^2 < \infty \tag{6.9}$$

Comme dans (6.4), les c_n sont les composantes de $|\varphi\rangle$. Vérifions que $|\varphi + \lambda\chi\rangle$ appartient à $\mathcal{H}$. Si $|\chi\rangle$ a pour composantes d_n, étant donné que

$$|c_n + \lambda d_n|^2 \leq 2(|c_n|^2 + |\lambda|^2 |d_n|^2)$$

il est clair que $||\varphi + \lambda\chi|| < \infty$. Le produit scalaire de deux vecteurs

$$\langle\chi|\varphi\rangle = \sum_{n=1}^{\infty} d_n^* c_n$$

est bien défini car, d'après l'inégalité de Schwartz (2.10)

$$|\langle\chi|\varphi\rangle| = \Big|\sum_{n=1}^{\infty} d_n^* c_n\Big| \leq ||\chi|| \ ||\varphi||$$

Vérifions ensuite que $\ell^{(2)}$ est complet. Soit $|\varphi^{(l)}\rangle$ et $|\varphi^{(m)}\rangle$ deux vecteurs de composantes $c_n^{(l)}$ et $c_n^{(m)}$. Si $||\varphi^{(m)} - \varphi^{(l)}|| < \varepsilon$ pour $l, m > N$, cela veut dire que

$$\left(\sum_{n=1}^{\infty} \left|c_n^{(l)} - c_n^{(m)}\right|^2\right)^{1/2} < \varepsilon$$

L'inégalité est *a fortiori* vraie pour chaque valeur individuelle de n et, pour n fixé, les nombres $c_n^{(l)}$ forment une suite de Cauchy qui converge vers c_n pour $l \to \infty$. On montre facilement (exercice 6.4.1) que le vecteur $\varphi^{(l)}$ converge vers $|\varphi\rangle = \sum_n c_n |n\rangle$ pour $l \to \infty$

$$\lim_{l\to\infty} \sum_n \left|c_n - c_n^{(l)}\right|^2 = \lim_{l\to\infty} ||\varphi - \varphi^{(l)}||^2 = 0$$

Enfin $\ell^{(2)}$ est de dimension dénombrable par construction.

(ii) *Espace* $L^{(2)}[a, b]$. Nous allons maintenant introduire une classe d'espaces vectoriels qui vont jouer un rôle capital, les espaces fonctionnels. L'exemple le plus simple est celui des fonctions de carré sommable sur l'intervalle $[a, b]$. Considérons les fonctions complexes $\varphi(x)$ telles que[4]

$$\int_a^b \mathrm{d}x |\varphi(x)|^2 < \infty \tag{6.10}$$

4. Deux fonctions $\varphi(x)$ et $\overline{\varphi}(x)$ telles que

$$\int_a^b \mathrm{d}x \, |\varphi(x) - \overline{\varphi}(x)|^2 = 0$$

représentent le même vecteur de $\mathcal{H}$: $||\varphi - \overline{\varphi}|| = 0$.

ou fonctions de carré sommable sur l'intervalle $[a, b]$. Ces fonctions forment un espace vectoriel, noté $L^{(2)}[a, b]$. En effet, (i) $\varphi(x) + \lambda\chi(x)$ est de carré sommable si $\varphi(x)$ et $\chi(x)$ le sont, (ii) le produit scalaire $\langle\chi|\varphi\rangle$

$$\langle\chi|\varphi\rangle = \int_a^b \mathrm{d}x\, \chi^*(x)\varphi(x) \tag{6.11}$$

est bien défini en raison de l'inégalité de Schwartz

$$\Big|\int_a^b \mathrm{d}x\, \chi^*(x)\varphi(x)\Big|^2 \le \int_a^b \mathrm{d}x\, |\chi(x)|^2 \int_a^b \mathrm{d}x\, |\varphi(x)|^2 = ||\chi||^2\, ||\varphi||^2 \tag{6.12}$$

Le fait que $L^{(2)}[a, b]$ soit complet résulte d'un théorème dû à Riesz et Fischer, et la séparabilité résulte d'un théorème standard de l'analyse de Fourier : toute fonction de carré sommable $\varphi(x)$ peut s'écrire, au sens de la convergence en moyenne (ou en norme), comme la somme d'une série de Fourier

$$\varphi(x) = \sum_{n=-\infty}^{\infty} c_n \frac{1}{\sqrt{(b-a)}} \exp\left(\frac{2\mathrm{i}\pi nx}{b-a}\right) \tag{6.13}$$

$$c_n = \frac{1}{\sqrt{(b-a)}} \int_a^b \mathrm{d}x\, \varphi(x) \exp\left(-\frac{2\mathrm{i}\pi nx}{b-a}\right) \tag{6.14}$$

Les fonctions

$$\varphi_n(x) = \frac{1}{\sqrt{(b-a)}} \exp\left(\frac{2\mathrm{i}\pi nx}{b-a}\right) \tag{6.15}$$

forment une base orthonormée dénombrable de $L^{(2)}[a, b]$, qui est donc un espace de Hilbert séparable.

(iii) *Espace* $L^{(2)}(\mathbb{R})$. Quand l'intervalle $[a, b]$ s'identifie à la droite réelle $\mathbb{R}$, $[a, b] \to]-\infty, +\infty[$, on obtient l'espace de Hilbert $L^{(2)}(\mathbb{R})$ (ou $L^{(2)}(]-\infty, +\infty[)$, l'espace des fonctions de carré sommable sur $]-\infty, +\infty[$. Bien que la démonstration soit plus délicate, on peut montrer que $L^{(2)}(\mathbb{R})$ reste un espace séparable, et donc isomorphe à $\ell^{(2)}$.

6.2 Opérateurs linéaires sur $\mathcal{H}$

6.2.1 Domaine et norme d'un opérateur

On définit des opérateurs linéaires sur $\mathcal{H}$ comme dans le cas de la dimension finie. Cependant, il existe des différences importantes. Il peut arriver, et c'est très souvent le cas en mécanique quantique, qu'un opérateur ne soit pas défini pour tout vecteur de $\mathcal{H}$, mais seulement sur un sous-ensemble de vecteurs de $\mathcal{H}$. Soit par exemple l'opérateur A agissant dans $\ell^{(2)}$ de la façon suivante : si $|\varphi\rangle$ a pour composantes $\{c_1, c_2, \ldots, c_n, \ldots\}$, alors $A|\varphi\rangle$ a pour composantes $\{c_1, 2c_2, \ldots, nc_n, \ldots\}$. Dans $L^{(2)}[a, b]$, cet opérateur correspond

à la différentiation à un facteur multiplicatif près, comme on le voit immédiatement en examinant la décomposition de Fourier (6.13). Il est clair que la norme au carré de $A|\varphi\rangle$, donnée par

$$||A\varphi||^2 = \sum_n n^2 |c_n|^2$$

peut diverger alors que $\sum_n |c_n|^2$ converge : il suffit par exemple de prendre $c_n = 1/n$. Autrement dit, $A|\varphi\rangle$ n'est pas nécessairement un vecteur de $\mathcal{H}$. On appelle *domaine* de A, noté $\mathcal{D}_A$, l'ensemble des vecteurs $|\varphi\rangle$ tel que $A|\varphi\rangle$ soit un vecteur de $\mathcal{H}$. Dans l'exemple ci-dessus, le domaine de A est l'ensemble des vecteurs tels que $\sum_n n^2|c_n|^2 < \infty$. Il est facile de se convaincre que ce domaine est dense dans $\mathcal{H}$. En pratique, un opérateur A ne présente un intérêt que si son domaine est dense dans $\mathcal{H}$.

Si $A|\varphi\rangle$ existe quel que soit $|\varphi\rangle$, on dit que l'opérateur A est *borné* : on doit alors avoir $||A\varphi|| < \infty$ quel que soit $|\varphi\rangle$. Le maximum de $||A\varphi||/||\varphi||$ est appelé la *norme* de A, qui est notée $||A||$

$$||A|| = \sup_{||\varphi||=1} ||A\varphi|| \tag{6.16}$$

Si la norme de $||A||$ n'existe pas, A est dit non borné. Les opérateurs non bornés sont d'un maniement beaucoup plus délicat que les opérateurs bornés. Malheureusement ils sont omniprésents en mécanique quantique.

Dans $L^{(2)}[0,1]$, l'opérateur X qui à $\varphi(x)$ fait correspondre la fonction $x\varphi(x)$

$$\varphi(x) \to (X\varphi)(x) = x\varphi(x) \tag{6.17}$$

est un opérateur borné de norme un. En revanche, l'opérateur $\mathrm{d}/\mathrm{d}x$, qui à $\varphi(x)$ fait correspondre sa dérivée

$$\varphi(x) \to \frac{\mathrm{d}\varphi(x)}{\mathrm{d}x} \tag{6.18}$$

n'est pas un opérateur borné. Nous l'avons déjà vu ci-dessus ; un autre argument simple consiste à trouver une fonction telle que la norme de $\varphi(x)$ soit finie, mais non celle de $\varphi'(x)$. Par exemple, si

$$\varphi(x) = x^{-1/4} \qquad \frac{\mathrm{d}\varphi(x)}{\mathrm{d}x} = -\frac{1}{4}\,x^{-5/4}$$

la norme de φ est finie, mais non celle de φ' car

$$\int_0^1 \mathrm{d}x\, x^{-1/2} = 2 \qquad \int_0^1 \mathrm{d}x\, \frac{1}{16}\, x^{-5/2} \ \text{ diverge à } x = 0$$

Les problèmes de domaine peuvent rendre délicats la définition de la somme et du produit de deux opérateurs non bornés. Par exemple, on ne peut *a*

priori définir la somme $A+B$ de deux opérateurs non bornés A et B que sur l'intersection $\mathcal{D}_A \cap \mathcal{D}_B$ des deux domaines, ce qui peut devenir problématique si cette intersection est réduite au vecteur nul ! Lorsque deux opérateurs A et B sont égaux sur un même domaine $\mathcal{D}_A$, mais que le domaine de B contient celui de A : $\mathcal{D}_A \subseteq \mathcal{D}_B$, on dit que B est un prolongement de A : $A \subseteq B$. Donnons un exemple : la relation de commutation canonique (4.45) entre les opérateurs position X et impulsion P, écrite pour une seule dimension d'espace ($d=1$)

$$[X, P] = \mathrm{i}\hbar I \tag{6.19}$$

implique qu'au moins un des deux opérateurs est non borné (exercice 6.4.3). Le membre de gauche $[X,P]$ de (6.19) n'est défini *a priori* que sur un sous-ensemble de $\mathcal{H}$, tandis que le membre de droite $\mathrm{i}\hbar I$ est défini pour tout vecteur de $\mathcal{H}$. L'écriture correcte de la relation de commutation canonique est donc

$$[X, P] \subseteq \mathrm{i}\hbar I$$

Notons une autre différence avec la dimension finie : alors que dans un espace vectoriel de dimension finie, l'existence d'un inverse à gauche entraîne celle d'un inverse à droite, et réciproquement, cette propriété n'est plus vraie en dimension infinie[5]. Soit par exemple les opérateurs A et B définis par leur action sur les composantes c_n d'un vecteur $|\varphi\rangle$

$$A(c_1, c_2, c_3 \ldots) = (c_2, c_3, c_4 \ldots) \qquad B(c_1, c_2, c_3 \ldots) = (0, c_1, c_2, \ldots)$$

alors

$$\begin{aligned} BA(c_1, c_2, c_3 \ldots) &= B(c_2, c_3, c_4 \ldots) = (0, c_2, c_3, \ldots) \\ AB(c_1, c_2, c_3 \ldots) &= A(0, c_1, c_2, \ldots) = (c_1, c_2, c_3, \ldots) \end{aligned}$$

et $AB = I$ tandis que $BA \neq I$, bien que A et B soient tous deux bornés.

6.2.2 Conjugaison hermitienne

Dans le cas d'un opérateur borné, il n'y a pas de difficulté de principe pour définir l'opérateur hermitien conjugué $A^\dagger$ de A par

$$\langle \chi | A\varphi \rangle = \langle A^\dagger \chi | \varphi \rangle \tag{6.20}$$

Comme dans le cas de la dimension finie, on dira que A est hermitien si $A = A^\dagger$ et on aura alors

$$\langle \chi | A\varphi \rangle = \langle A\chi | \varphi \rangle \tag{6.21}$$

Les choses se compliquent si A n'est pas borné en raison des questions de domaine. Tout d'abord, (6.20) ne peut définir $A^\dagger$ que si $\mathcal{D}_A$ est dense dans $\mathcal{H}$.

5. Un exemple important d'un tel opérateur en physique est l'opérateur de Møller de la théorie de la diffusion : exercice 13.6.9.

Ensuite, le domaine de définition de $A^\dagger$ est en général plus grand que celui de $A : \mathcal{D}_A \subseteq \mathcal{D}_{A^\dagger}$. Nous allons le voir sur un exemple dans un instant. En général, pour un opérateur non borné vérifiant (6.21), on n'aura pas $A = A^\dagger$ mais plutôt $A \subseteq A^\dagger$. Les mathématiciens réservent la dénomination “opérateurs hermitiens” aux opérateurs tels que $A \subseteq A^\dagger$, et appellent “auto-adjoints” les opérateurs tels que $A = A^\dagger$.

Illustrons cette discussion par un exemple dans $L^{(2)}[0,1]$, qui va nous familiariser avec le produit scalaire et la conjugaison hermitienne dans cet espace. Soit A_0 l'opérateur $-\mathrm{i}\,\mathrm{d}/\mathrm{d}x$, défini sur le domaine $\mathcal{D}_{A_0}$ des fonctions $\varphi(x)$ de $L^{(2)}[0,1]$, dérivables et dont la dérivée est de carré sommable, et vérifiant de plus les conditions aux limites $\varphi(0) = \varphi(1) = 0$, d'où l'indice 0 de A_0. Il est intuitivement évident, et facile à vérifier, que ce domaine est dense dans $L^{(2)}[0,1]$. Montrons d'abord que A_0 est hermitien ; $\chi(x)$ étant une fonction de $L^{(2)}[0,1]$ dérivable et dont la dérivée appartient à $L^{(2)}[0,1]$

$$\begin{aligned}
\langle\chi|A_0\varphi\rangle &= \int_0^1 \mathrm{d}x\,\chi^*(x)\Big(-\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}\,\varphi(x)\Big) = -\mathrm{i}\int_0^1 \mathrm{d}x\,\chi^*(x)\varphi'(x) \\
\langle A_0\chi|\varphi\rangle &= \int_0^1 \mathrm{d}x\,\Big(-\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}\,\chi(x)\Big)^*\varphi(x) = \mathrm{i}\int_0^1 \mathrm{d}x\,(\chi'(x))^*\varphi(x) \\
\langle\chi|A_0\varphi\rangle - \langle A_0\chi|\varphi\rangle &= -\mathrm{i}[\chi^*(x)\varphi(x)]_0^1 = 0 \qquad (6.22)
\end{aligned}$$

On remarquera la nécessité pour l'hermiticité du facteur i et des conditions aux limites. On peut définir $A_0^\dagger$ sur un domaine plus grand que $\mathcal{D}_{A_0}$. En effet, pour des fonctions $\chi(x)$ non contraintes par des conditions aux limites, c'est-à-dire telles que $\chi(0)$ et $\chi(1)$ soient quelconques

$$\begin{aligned}
\langle A_0^\dagger\chi|\varphi\rangle &= \mathrm{i}\int_0^1 \mathrm{d}x\,(\chi'(x))^*\varphi(x) \\
&= \mathrm{i}[\chi^*(x)\varphi(x)]_0^1 - \mathrm{i}\int_0^1 \mathrm{d}x\,\chi^*(x)\varphi'(x) = \langle\chi|A_0\varphi\rangle
\end{aligned}$$

et par conséquent $A_0 \subseteq A_0^\dagger$. Enfin, définissons A_C comme l'opérateur $-\mathrm{i}\,\mathrm{d}/\mathrm{d}x$ agissant dans le domaine $\mathcal{D}_{A_C}$ des fonctions $\varphi(x)$ de $L^{(2)}[0,1]$, dérivables, dont la dérivée appartient à $L^{(2)}[0,1]$, et vérifiant les conditions aux limites

$$\varphi(1) = C\varphi(0) \qquad |C| = 1$$

L'opérateur A_C est auto-adjoint. En effet,

$$\langle A_C\chi|\varphi\rangle - \langle\chi|A_C\varphi\rangle = -\mathrm{i}(C\chi^*(1) - \chi^*(0))\varphi(0)$$

La condition nécessaire et suffisante pour que le membre de droite s'annule[6] est que $\chi(1) = C\chi(0)$, ce qui montre que le domaine de l'opérateur hermitien

6. Noter que $C^* = 1/C$.

conjugué est aussi $\mathcal{D}_{A_C}$: $A_C^\dagger = A_C$. Les opérateurs A_C représentent pour chaque valeur de C des *prolongements différents* de A_0 : même si la définition est superficiellement la même ($A = -\mathrm{i}\mathrm{d}/\mathrm{d}x$), la différence des domaines fait que A_C et $A_{C'}$ sont des opérateurs différents pour $C \neq C'$! On le vérifie en montrant que les valeurs propres et vecteurs propres de A_C et $A_{C'}$ sont différents pour $C \neq C'$ (exercice 6.4.3).

6.3 Décomposition spectrale

6.3.1 Opérateurs hermitiens

Le théorème de décomposition spectrale qui généralise (2.31) est en toute rigueur valable uniquement pour les opérateurs auto-adjoints[7]. Suivant la tradition des physiciens, nous ne ferons désormais plus la différence entre hermitien et auto-adjoint, et nous parlerons uniquement d'opérateurs hermitiens. Si un opérateur A est hermitien, et même s'il est borné, l'équation aux valeurs propres

$$A|\varphi\rangle = a|\varphi\rangle \tag{6.23}$$

n'a pas toujours de solution. Par exemple, dans $L^{(2)}(\mathbb{R})$ l'opérateur $-\mathrm{i}\mathrm{d}/\mathrm{d}x$ est hermitien, ce que l'on voit par une généralisation immédiate de (6.22). L'équation

$$-\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}\,\varphi(x) = a\varphi(x) \tag{6.24}$$

a pour solution l'onde plane

$$\varphi_a(x) = C\,\mathrm{e}^{\,\mathrm{i}ax} \tag{6.25}$$

où C est une constante, mais $\varphi_a(x)$ n'appartient pas à $L^{(2)}(\mathbb{R})$ car

$$\int_{-\infty}^{\infty} \mathrm{d}x\,|\varphi_a(x)|^2 = \int_{-\infty}^{\infty} \mathrm{d}x\,|C|^2$$

est une intégrale divergente ; $-\mathrm{i}\mathrm{d}/\mathrm{d}x$ est un opérateur non borné, mais même pour un opérateur borné, par exemple la multiplication par x dans $L^{(2)}[0,1]$, l'équation

$$x\chi_a(x) = a\chi_a(x) \tag{6.26}$$

n'a pas de solution dans $L^{(2)}[0,1]$. En fait, la généralisation de (6.23) au cas de la dimension infinie n'est assurée que pour une classe très particulière d'opérateurs, les opérateurs compacts.

En dimension finie, lorsque $|\varphi\rangle$ est vecteur propre de A avec la valeur propre a suivant (6.23), on dit que a appartient au *spectre* de A. Pour généraliser cette notion à la dimension infinie, considérons l'opérateur $(zI - A)$, où

---

7. Plus précisément, pour les opérateurs “essentiellement auto-adjoints” : $(A^\dagger)^\dagger = A^\dagger$.

z est un nombre complexe et l'équation

$$(zI - A)|\varphi\rangle = |\chi\rangle \tag{6.27}$$

Soit $\mathcal{D}$ le domaine de $(zI - A)$ et $\Delta(z)$ son image. Si $\Delta(z) = \mathcal{H}$, z est une valeur régulière de A : la correspondance entre $|\varphi\rangle$ et $|\chi\rangle$ est biunivoque et la résolvante (2.46) $R(z, A) = (zI - A)^{-1}$ existe. Le spectre de A est par définition l'ensemble des valeurs de z non régulières. Cette définition coïncide bien avec celle de la dimension finie : en effet si $|\varphi\rangle$ vérifie (6.23)

$$(zI - A)\Big|_{z=a}|\varphi\rangle = (aI - A)|\varphi\rangle = 0$$

et la résolvante n'est pas définie pour $z = a$. Si A est hermitien, il est facile de montrer (exercice 6.4.2) que $z = a + \mathrm{i}b$ est une valeur régulière lorsque $b \neq 0$: le spectre de A est donc réel, comme dans le cas de la dimension finie. Les valeurs de a peuvent, soit être indicées par un indice discret : $a_1, a_2, \ \ldots, a_n, \ \ldots$, soit prendre des valeurs continues, par exemple toutes les valeurs sur un intervalle de la droite réelle : on distingue donc un *spectre discret* et un *spectre continu*. Les valeurs de a, appartenant au spectre discret, vérifient une équation aux valeurs propres (6.23), mais non celles du spectre continu. Le spectre continu et le spectre discret peuvent se recouvrir : par exemple si a prend toute les valeurs entre 0 et 1, il peut arriver que le spectre de A contienne des valeurs propres discrètes $0 \leq a_n \leq 1$, bien que ce cas soit exceptionnel en pratique. En général, pour les opérateurs utilisés en physique quantique, spectre discret et continu ne se recouvrent pas.

Bien que le spectre de la dimension infinie présente des propriétés nouvelles par rapport à celui de la dimension finie, il existe un théorème de décomposition spectrale qui généralise (2.31)

$$A = \sum_n a_n \mathcal{P}_n$$

La forme mathématique précise de ce théorème est complexe, et les physiciens s'en sortent en utilisant des “pseudo-vecteurs propres”, c'est-à-dire comme dans (6.25) des objets qui vérifient formellement l'équation aux valeurs propres mais ne sont pas des éléments de $\mathcal{H}$. Dans le cas de (6.26), la “solution” sera

$$\chi_a(x) = \delta(x - a) \quad \text{car} \quad x\delta(x - a) = a\delta(x - a) \tag{6.28}$$

où $\delta(x)$ est la distribution de Dirac, qui n'est pas une fonction, et certainement pas un élément de $L^{(2)}[0, 1]$.

Les exemples que nous venons de donner nous mettent sur la voie du résultat général. La condition de “normalisation” des pseudo-vecteurs propres (6.25) de $-\mathrm{i}\mathrm{d}/\mathrm{d}x$ est, avec le choix $C = 1/\sqrt{2\pi}$

$$\langle\varphi_a|\varphi_b\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}x \, \mathrm{e}^{-\mathrm{i}ax}\mathrm{e}^{ibx} = \delta(a - b) \tag{6.29}$$

tandis que pour les valeurs propres (6.28) de x

$$\langle \chi_a | \chi_b \rangle = \int_{-\infty}^{\infty} \mathrm{d}x\, \delta(x-a)\delta(x-b) = \delta(a-b) \tag{6.30}$$

La normalisation des pseudo-vecteurs propres est donc donnée, non par un delta de Kronecker, mais par un delta de Dirac. La généralisation du théorème de décomposition spectrale s'énonce ainsi.

- Pour les valeurs a_n du spectre discret étiquetées par un indice discret n, on peut écrire une équation aux valeurs propres et des conditions de normalisation analogues à celles de la dimension finie

$$A|n,r\rangle = a_n|n,r\rangle \tag{6.31}$$
$$\langle n,r|n',r'\rangle = \delta_{nn'}\,\delta_{rr'} \tag{6.32}$$

  où r est un indice de dégénérescence discret.
- Pour les valeurs $a(\nu)$ du spectre continu étiquetées par un indice continu ν, nous aurons

$$A|\nu,s\rangle = a(\nu)|\nu,s\rangle \tag{6.33}$$
$$\langle \nu,s|\nu',s'\rangle = \delta(\nu-\nu')\,\delta_{ss'} \tag{6.34}$$

  où $|\nu,s\rangle$ n'est pas un vecteur de $\mathcal{H}$; s est un indice de dégénérescence qui peut être discret ou continu, mais que nous avons pris discret pour fixer les notations.
- En outre, les vecteurs propres du spectre discret et ceux du spectre continu sont orthogonaux

$$\langle n,r|\nu,s\rangle = 0 \tag{6.35}$$

La généralisation de la décomposition de l'identité, ou relation de fermeture (2.30) s'écrit

$$\boxed{I = \sum_{n,r} |n,r\rangle\langle n,r| + \sum_s \int \mathrm{d}\nu\, |\nu,s\rangle\langle \nu,s|} \tag{6.36}$$

tandis que la décomposition spectrale (2.31) de A devient

$$\boxed{A = \sum_{n,r} |n,r\rangle a_n \langle n,r| + \sum_s \int \mathrm{d}\nu\, |\nu,s\rangle a(\nu) \langle \nu,s|} \tag{6.37}$$

Insistons sur le fait que l'existence d'un spectre discret et/ou continu n'est en rien liée au fait que l'opérateur A soit ou non borné : il existe des opérateurs non bornés dont le spectre est entièrement discret, comme le hamiltonien de l'oscillateur harmonique (§ 10.1.1) ou le carré du moment angulaire $\vec{J}^2$ (section 9.1), et des opérateurs bornés comme la multiplication par x dans $L^{(2)}[0,1]$ dont le spectre est entièrement continu.

6.3.2 Opérateurs unitaires

Un opérateur unitaire est défini par

$$U^\dagger U = UU^\dagger = I \quad \text{ou} \quad U^\dagger = U^{-1} \tag{6.38}$$

Comme dans le cas de la dimension finie, on peut construire des opérateurs unitaires par exponentiation d'opérateurs hermitiens. Utilisant la décomposition spectrale de A (6.37)

$$\boxed{U(\alpha)=\exp(\mathrm{i}\alpha A)=\sum_{n,r}|n,r\rangle\exp(\mathrm{i}\alpha a_n)\langle n,r|+\sum_s\int \mathrm{d}\nu\,|\nu,s\rangle\exp[\mathrm{i}\alpha a(\nu)]\langle \nu,s|} \tag{6.39}$$

Cette équation montre que le spectre de $\exp(\mathrm{i}\alpha A)$ est localisé sur le cercle $|z|=1$, et il est facile de vérifier que cette propriété est vraie de tout opérateur unitaire. De plus, (6.39) montre que $U(\alpha)$ vérifie la propriété de groupe abélien

$$U(\alpha_1+\alpha_2)=U(\alpha_1)U(\alpha_2) \qquad U(0)=I \tag{6.40}$$

La réciproque de cette propriété est un théorème important, le théorème de Stone[8].

Théorème de Stone. Soit un ensemble d'opérateurs unitaires dépendant d'un paramètre continu α et vérifiant la loi de groupe abélien (6.40). Il existe alors un opérateur hermitien T, appelé *générateur infinitésimal* du groupe de transformations $U(\alpha)$ tel que $U(\alpha)=\exp(\mathrm{i}\alpha T)$.■

On peut donner une démonstration heuristique de ce théorème, en montrant que $U(\alpha)$ vérifie une équation différentielle. Si $\delta\alpha \to 0$

$$U(\alpha+\delta\alpha)=U(\delta\alpha)U(\alpha)\simeq\left(I+\delta\alpha\frac{\mathrm{d}U}{\mathrm{d}\alpha}\Big|_{\alpha=0}\right)U(\alpha) \tag{6.41}$$

Si l'on pose

$$T=-\mathrm{i}\,\frac{\mathrm{d}U}{\mathrm{d}\alpha}\Big|_{\alpha=0} \tag{6.42}$$

T doit être hermitien car

$$\begin{aligned} U(\delta\alpha)U^\dagger(\delta\alpha) &\simeq (I+\mathrm{i}\,\delta\alpha\,T)\left(I-\mathrm{i}\,\delta\alpha\,T^\dagger\right) \\ &\simeq I+\mathrm{i}\,\delta\alpha(T-T^\dagger)=I \end{aligned}$$

d'où $T=T^\dagger$. On déduit de (6.41)

$$\frac{\mathrm{d}U(\alpha)}{\mathrm{d}\alpha}=\mathrm{i}TU(\alpha) \tag{6.43}$$

ce qui donne le théorème de Stone par intégration et en tenant compte de la condition $U(0)=I$.

8. Aussi appelé théorème SNAG : Stone, Naimark, Ambrose et Godement.

6.4 Exercices

6.4.1 Espaces de dimension infinie

1. Montrer que l'espace ℓ^2 est complet.

2. Montrer que la convergence forte implique la convergence faible, mais non l'inverse, sauf si l'espace est de dimension finie.

6.4.2 Spectre d'un opérateur hermitien

Montrer que si $A = A^\dagger$ et $z = x + \mathrm{i}y$, le vecteur

$$|\chi\rangle = (zI - A)|\varphi\rangle$$

ne peut pas s'annuler si $y \neq 0$.

6.4.3 Relations de commutation canoniques

1. Soit deux opérateurs hermitiens A et B vérifiant la relation de commutation $[B, A] = \mathrm{i}I$. Montrer que l'un au moins des deux opérateurs est non borné. On pourra supposer sans restreindre la généralité (pourquoi ?) que $||B|| = 1$. Suggestion : montrer que

$$[B, A^n] = \mathrm{i}nA^{n-1}$$

et en déduire

$$||A^n|| \geq \frac{n}{2}\,||A^{n-1}||$$

2. On suppose que A possède un vecteur propre normalisable $|\varphi\rangle$

$$A|\varphi\rangle = a|\varphi\rangle \qquad a = a^*$$

On a, d'une part,

$$\begin{aligned}\langle\varphi|(BA - AB)|\varphi\rangle &= \langle\varphi|B|A\varphi\rangle - \langle A\varphi|B|\varphi\rangle \\ &= a(\langle\varphi|B|\varphi\rangle - \langle\varphi|B|\varphi\rangle) = 0\end{aligned}$$

et, d'autre part,

$$\langle\varphi|(BA - AB)|\varphi\rangle = \langle\varphi|[B, A]|\varphi\rangle = \mathrm{i}||\varphi||^2$$

Quelle est la solution de ce pseudo-paradoxe ? Suggestion : examiner le cas où $B = X$ (multiplication par x) et $A = -\mathrm{i}\,\mathrm{d}/\mathrm{d}x$ sur $L^2[0, 1]$ avec les conditions aux limites $\varphi(x = 0) = \varphi(x = 1)$.

3. On considère les opérateurs A_C définis au § 6.2.2. Trouver les valeurs propres et les vecteurs propres de A_C, et montrer que le spectre de A_C est différent suivant les valeurs de C. Le théorème de von Neumann (chapitre 7)

énonce que les relations de commutation canoniques sont uniques à une équivalence unitaire près. Pourtant,

$$[X, A_C] = \mathrm{i}I \quad \text{et} \quad [X, A_{C'}] = \mathrm{i}I$$

et $A_C \neq A_{C'}$ si $C \neq C'$. Quelle est la solution de ce nouveau pseudo-paradoxe (non indépendant du précédent) ?

6.4.4 Opérateurs de dilatation et de transformation conforme

1. Soit A l'opérateur

$$A = -\mathrm{i}\, x \frac{\partial}{\partial x}$$

A est-il hermitien ? Montrer que

$$\left[\mathrm{e}^{-\mathrm{i}\alpha A}\, \Phi\right](x) = \Phi(\mathrm{e}^{-\alpha}\, x)$$

Méthode 1 : utiliser la variable $u = \ln x$

2. Méthode 2 : obtenir l'équation aux dérivées partielles

$$\left(\frac{\partial}{\partial \alpha} + x \frac{\partial}{\partial x}\right) \left[\mathrm{e}^{-\mathrm{i}\alpha A}\, \Phi\right](x) = 0$$

3. Soit B l'opérateur

$$B = -\mathrm{i}\, x^2 \frac{\partial}{\partial x}$$

Montrer que

$$\left[\mathrm{e}^{-\mathrm{i}\alpha B}\, \Phi\right](x) = \Phi\left(\frac{x}{1 + \alpha x}\right)$$

6.5 Bibliographie

Jauch [1968], chapitres 1 à 4 et Peres [1993], chapitre 4, contiennent un exposé assez détaillé et mathématiquement rigoureux des notions utiles sur les espaces de Hilbert de dimension infinie et les opérateurs sur ces espaces. Le lecteur porté sur les aspects mathématiques pourra se plonger dans le livre classique de Riesz et Nagy [1955].

Chapitre 7

Symétries en physique quantique

La résolution de problèmes de physique classique se simplifie, parfois de façon considérable, en présence de *symétries*, c'est-à-dire de transformations qui laissent invariantes certaines propriétés physiques. Par exemple, en mécanique classique, le problème d'une particule dans une force centrale $\vec{F} = F(r)\hat{r}$ indépendante du temps est invariant par translation de temps et par rotation autour de tout axe passant par l'origine. L'invariance par translation de temps assure la conservation de l'énergie mécanique E, et l'invariance par rotation la conservation du moment angulaire $\vec{J}$. En l'absence de symétries, on doit *a priori* résoudre un système de trois équations différentielles du second ordre (une par composante). Grâce à ces symétries, on se ramène à la résolution d'une seule équation différentielle du premier ordre. Résumons ci-dessous les conséquences des principales invariances en mécanique classique.

- L'invariance par translation de temps de l'énergie potentielle V entraîne la conservation de l'énergie mécanique $K + V$, somme de l'énergie cinétique K et de l'énergie potentielle V.
- L'invariance de l'énergie potentielle par translation d'espace parallèle à un vecteur $\hat{n}$ entraîne la conservation de la composante $\vec{P} \cdot \hat{n} = P_{\hat{n}}$ de l'impulsion.
- L'invariance de l'énergie potentielle par rotation autour d'un axe $\hat{n}$ entraîne la conservation de la composante $\vec{J} \cdot \hat{n} = J_{\hat{n}}$ du moment angulaire.

Les propriétés de symétrie jouent un rôle encore plus important en mécanique quantique. Elles permettent d'obtenir des résultats très généraux, qui sont indépendants des approximations faites par exemple pour le hamiltonien (bien sûr si ces approximations respectent les symétries du problème !).

Dans ce chapitre, nous exploiterons les hypothèses d'invariance suivantes, que nous supposerons valables[1] pour un système isolé.

- La description d'un système isolé ne doit pas dépendre de l'origine des temps : elle doit être invariante par translation de l'origine des temps.
- L'espace est homogène, ce qui veut dire que la description d'un système isolé ne doit pas dépendre de l'origine des axes : elle doit être invariante par translation d'espace.
- L'espace est isotrope, ce qui veut dire que la description d'un système isolé ne doit pas dépendre de l'orientation choisie pour les axes : elle doit être invariante par rotation.
- La forme des lois physiques doit être inchangée lorsque l'on passe d'un référentiel d'inertie à un autre.

Cette dernière hypothèse doit être précisée, car il existe deux lois de transformation possibles entre référentiels d'inertie, la transformation de Lorentz et celle de Galilée, cette dernière étant valable lorsque $v/c \to 0$, où v est une vitesse caractéristique du problème concerné. Naturellement, c'est la transformation de Lorentz que l'on doit choisir en général, mais on ne peut alors éviter le cadre de la théorie quantique des champs. À l'exception du chapitre 19, nous considérerons uniquement des particules massives dont les vitesses sont faibles par rapport à la vitesse de la lumière[2] et nous pourrons donc nous limiter à la transformation de Galilée et travailler dans le cadre de ce qui est appelé conventionnellement, mais improprement[3], la “mécanique quantique non relativiste”.

7.1 Transformation d'un état dans une opération de symétrie

7.1.1 Invariance des probabilités dans une opération de symétrie

Le point de vue adopté implicitement dans l'introduction de ce chapitre était le point de vue dit *passif* : le système physique est inchangé, mais on

1. Ces hypothèses sont éminemment plausibles, mais, après tout, il pourrait exister des effets subtils qui remettent en cause une (ou plusieurs) de ces invariances. Avant 1957, l'immense majorité des physiciens auraient parié sur l'invariance de la physique par l'opération parité. Pauli avait même interdit que l'on fasse au CERN à Genève une expérience destinée à montrer l'éventuelle violation de cette invariance, tellement il trouvait cette possibilité absurde. En conséquence, la violation de l'invariance par parité fut découverte aux USA dans l'expérience de C.S. Wu (*cf.* 7.3.3).

2. Cependant, nous prendrons en compte les interactions de ces particules avec des photons, mais sans entrer dans les complications de l'électrodynamique quantique relativiste.

3. En effet, cette théorie est parfaitement relativiste, puisqu'elle obéit à la relativité. . .galiléenne !

modifie le système d'axes. Il est en général équivalent[4] d'adopter le point de vue *actif*, où le système d'axes est inchangé, et où on applique une opération de symétrie sur le système physique. Nous avions d'ailleurs déjà utilisé cette équivalence dans la discussion du § 3.2.4. Dans la suite de ce chapitre, nous allons adopter le point de vue actif, qui est peut-être plus intuitif (au moins pour l'auteur !) et sera plus commode pour certaines discussions, par exemple celles de la section 9.5.

Nous avons vu au chapitre 4, postulat **I**, que l'objet mathématique en correspondance biunivoque avec un état physique était un rayon unitaire de l'espace des états $\mathcal{H}$, c'est-à-dire un vecteur unitaire à un facteur de phase près. Dans cette section *uniquement*, la distinction entre vecteurs et rayons sera cruciale ; nous pourrons l'oublier par la suite. On vérifie immédiatement que la relation entre deux vecteurs de $\mathcal{H}$

$$|\varphi'\rangle = \mathrm{e}^{\,\mathrm{i}\theta}|\varphi\rangle \tag{7.1}$$

où θ est un nombre réel, est une relation d'équivalence[5] $|\varphi'\rangle \sim |\varphi\rangle$. La classe d'équivalence est un rayon, que nous noterons $\tilde{\varphi}$. Le produit scalaire de deux rayons $\tilde{\varphi}$ et $\tilde{\chi}$ n'est pas défini, mais le module de ce produit scalaire, que nous noterons $|(\tilde{\chi}, \tilde{\varphi})|$ est bien défini : on peut choisir deux représentants arbitraires $|\varphi\rangle$ et $|\chi\rangle$ dans les classes d'équivalence et écrire

$$|(\tilde{\chi}, \tilde{\varphi})| = |\langle\chi|\varphi\rangle| \tag{7.2}$$

car les facteurs de phase disparaissent lorsque l'on prend le module. Le résultat est indépendant du choix des représentants dans les classes d'équivalence.

Revenons au spin 1/2 du chapitre 3 : nous avons vu comment préparer un état de spin orienté suivant Oz que nous représenterons par le rayon $\tilde{\varphi}_+$, en utilisant un appareil de Stern-Gerlach dont le champ magnétique est orienté suivant Oz et en sélectionnant les atomes déviés vers le haut (en choisissant un signe approprié pour le champ). Faisons tourner le champ d'un angle α autour de la direction de propagation Oy pour l'amener suivant une direction $\hat{n}_\alpha$ faisant un angle α avec Oz, $0 \le \alpha < 2\pi$. Nous préparons ainsi l'état représenté par le rayon $\tilde{\varphi}_+(\hat{n}_\alpha)$, qui sera par définition l'état transformé de $\tilde{\varphi}_+$ par une rotation de α autour de Oy (figure 7.1). Avec les notations du chapitre 3, la classe d'équivalence du vecteur $|+\rangle$ est le rayon $\tilde{\varphi}_+$, celle du vecteur $|+, \hat{n}_\alpha\rangle$, le rayon $\tilde{\varphi}_+(\hat{n}_\alpha)$. En général, le transformé $\tilde{\varphi}_{\mathcal{R}}$ par une rotation $\mathcal{R}$ d'un état $\tilde{\varphi}$ sera obtenu en effectuant une rotation $\mathcal{R}$ sur l'appareil qui prépare $\tilde{\varphi}$.

Supposons maintenant qu'à la suite du premier appareil de Stern-Gerlach dont le champ est parallèle à Oz, le polariseur, on place un second appareil, l'analyseur, dont le champ est parallèle à la direction $\hat{n}_\beta$ obtenue à partir de

4. Pour certaines transformations comme la réflexion par rapport à un plan, il est plus simple d'utiliser le point de vue passif, qui consiste à regarder le système dans un miroir, mais on peut aussi imaginer de construire un appareillage symétrique de l'original par rapport à un plan.

5. La notation $\sim$ désigne ici une relation d'équivalence, et non "de l'ordre de".

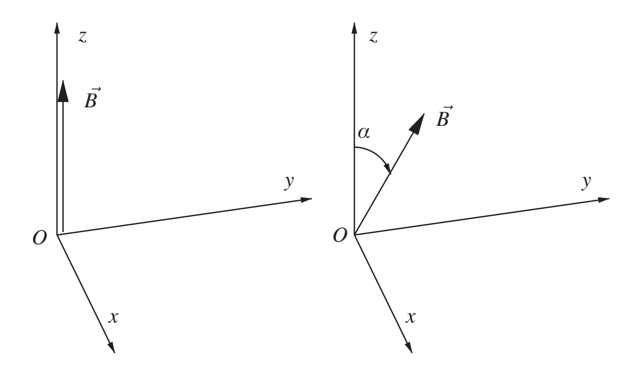


FIG. 7.1 – Préparation des états (rayons) $\tilde{\varphi}_+$ et $\tilde{\varphi}_+(\hat{n}_\alpha)$.

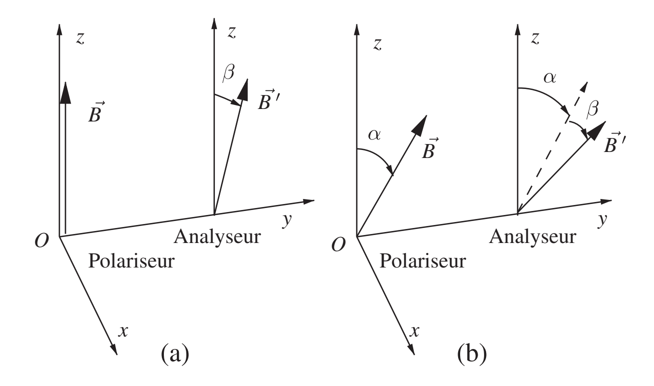


FIG. 7.2 – Rotations simultanées du polariseur et de l'analyseur d'un angle α.

Oz par une rotation d'angle β autour de Oy (figure 7.2a). S'il n'y a pas sur le trajet de champ magnétique susceptible de faire tourner le spin, la probabilité pour que le spin soit dévié dans la direction $\hat{n}_\beta$ est

$$|(\tilde{\varphi}_+(\hat{n}_\beta), \tilde{\varphi}_+)|^2$$

Effectuons maintenant l'expérience en faisant tourner *à la fois* le polariseur et l'analyseur d'un angle α (figure 7.2b). La probabilité de déviation dans la direction $\hat{n}_{\alpha+\beta}$ est

$$|(\tilde{\varphi}_+(\hat{n}_{\alpha+\beta}), \tilde{\varphi}_+(\hat{n}_\alpha))|^2$$

Comme on a fait subir la même rotation au polariseur et à l'analyseur, l'invariance par rotation implique que les probabilités sont inchangées

$$|(\tilde{\varphi}_+(\hat{n}_{\alpha+\beta}), \tilde{\varphi}_+(\hat{n}_\alpha))|^2 = |(\tilde{\varphi}_+(\hat{n}_\beta), \tilde{\varphi}_+)|^2 \tag{7.3}$$

Généralisons (7.3) : si l'on effectue une transformation g sur un état $\tilde{\varphi}$ en appliquant cette transformation sur l'appareil de préparation de $\tilde{\varphi}$ pour obtenir l'état transformé $\tilde{\varphi}_g$: $\tilde{\varphi} \to \tilde{\varphi}_g$, et si l'on effectue la même opération

sur l'appareil de mesure pour $\tilde{\chi}$: $\tilde{\chi} \to \tilde{\chi}_g$ alors les probabilités doivent être inchangées si la physique est invariante dans cette opération

$$|(\tilde{\chi}_g, \tilde{\varphi}_g)|^2 = |(\tilde{\chi}, \tilde{\varphi})|^2 \tag{7.4}$$

7.1.2 Théorème de Wigner

La propriété (7.4) sur les *rayons* se traduit par une propriété sur les *vecteurs* grâce à un théorème d'une grande importance dû à Wigner.

Théorème de Wigner. Si l'on traduit mathématiquement la loi de transformation des états physiques par une loi de transformation sur les *rayons* correspondants : $\tilde{\varphi} \to \tilde{\varphi}_g$ lorsque l'on applique une transformation g à un système physique, et si l'on suppose que les probabilités sont invariantes dans cette transformation

$$|(\tilde{\chi}_g, \tilde{\varphi}_g)|^2 = |(\tilde{\chi}, \tilde{\varphi})|^2 \qquad \forall\, \tilde{\varphi},\, \tilde{\chi}$$

alors il est possible de choisir un représentant $|\varphi_g\rangle$ de $\tilde{\varphi}_g$ tel que pour tout vecteur $|\varphi\rangle \in \mathcal{H}$

$$|\varphi_g\rangle = U(g)|\varphi\rangle \tag{7.5}$$

où l'opérateur $U(g)$ est unitaire ou antiunitaire et est unique à un facteur de phase près.■

La loi de transformation des *rayons* devient donc une loi de transformation des *vecteurs*, par application d'un opérateur qui ne dépend que de la transformation g. Si $U(g)$ est unitaire, le théorème de Wigner implique non seulement l'invariance de la norme du produit scalaire, mais aussi celle de sa phase, puisque

$$\langle U(g)\chi | U(g)\varphi\rangle = \langle \chi|\varphi\rangle$$

Les opérateurs antiunitaires transforment le produit scalaire en son complexe conjugué

$$\langle U(g)\chi | U(g)\varphi\rangle = \langle \chi|\varphi\rangle^* = \langle \varphi|\chi\rangle \tag{7.6}$$

La démonstration du théorème de Wigner ne fait intervenir que des notions élémentaires, mais elle est fastidieuse, et nous la renvoyons à l'appendice A1. Les opérateurs antiunitaires n'interviennent que lorsque la transformation g inclut le renversement du sens du temps ; nous en dirons un mot au § 7.3.3, mais nous en renvoyons l'étude détaillée à l'appendice A. Nous nous limitons désormais aux transformations unitaires.

Le théorème de Wigner a des conséquences particulièrement intéressantes si les transformations g forment un groupe $\mathcal{G}$. Le produit $g = g_2 g_1$ de deux transformations, de même que la transformation inverse g^{-1}, sont alors des transformations de $\mathcal{G}$. L'ordre des transformations dans $g_2 g_1$ est important car le groupe $\mathcal{G}$ n'est pas en général abélien : $g_2 g_1 \neq g_1 g_2$. Si $g = g_2 g_1$, les rayons $\tilde{\varphi}_g$ et $\tilde{\varphi}_{g_2 g_1}$ doivent être identiques. Par exemple, si $\mathcal{G}$ est le groupe des rotations autour de Oz, et si $\mathcal{R}_z(\theta)$ représente la rotation d'angle θ autour de Oz, on a

$$\mathcal{R}_z(\theta = \theta_2 + \theta_1) = \mathcal{R}_z(\theta_2)\mathcal{R}_z(\theta_1) \tag{7.7}$$

L'état physique obtenu en effectuant une rotation d'angle $\theta = \theta_2 + \theta_1$ doit être identique à celui obtenu en effectuant d'abord une rotation d'angle θ_1 suivie d'une rotation d'angle θ_2.

Utilisons maintenant le théorème de Wigner pour faire un choix de phases sur les vecteurs tel que la correspondance entre $|\varphi\rangle$ et $|\varphi_g\rangle$ soit donnée par (7.5). Nous avons, d'une part,

$$|\varphi_g\rangle = U(g)|\varphi\rangle \tag{7.8}$$

et, d'autre part,

$$|\varphi_{g_2 g_1}\rangle = U(g_2)|\varphi_{g_1}\rangle = U(g_2)U(g_1)|\varphi\rangle \tag{7.9}$$

Les vecteurs $|\varphi_g\rangle$ et $|\varphi_{g_2 g_1}\rangle$ représentent des états physiques identiques, et ils doivent être égaux à un facteur de phase près

$$|\varphi_g\rangle = \mathrm{e}^{\,\mathrm{i}\alpha(g_2, g_1)}|\varphi_{g_2 g_1}\rangle \tag{7.10}$$

Le facteur de phase dans (7.10) pourrait *a priori* dépendre de $|\varphi\rangle$, mais en fait, il dépend uniquement de g_1 et de g_2. En effet, si nous écrivons

$$|\varphi_g\rangle = \mathrm{e}^{\,\mathrm{i}\alpha}|\varphi_{g_2 g_1}\rangle \qquad |\chi_g\rangle = \mathrm{e}^{\,\mathrm{i}\beta}|\chi_{g_2 g_1}\rangle$$

nous pouvons examiner le produit scalaire $\langle\chi|\varphi\rangle$

$$\begin{aligned}\langle\chi|\varphi\rangle &= \langle\chi_g|\varphi_g\rangle = \mathrm{e}^{\,\mathrm{i}(\alpha-\beta)}\langle\chi_{g_2 g_1}|\varphi_{g_2 g_1}\rangle \\ &= \mathrm{e}^{\,\mathrm{i}(\alpha-\beta)}\langle U(g_2)U(g_1)\chi|U(g_2)U(g_1)\varphi\rangle \\ &= \mathrm{e}^{\,\mathrm{i}(\alpha-\beta)}\langle\chi|\varphi\rangle\end{aligned}$$

ce qui implique $\alpha = \beta$. Comme le vecteur $|\varphi\rangle$ est arbitraire, (7.10) entraîne une relation correspondante pour les opérateurs $U(g)$

$$U(g) = \mathrm{e}^{\mathrm{i}\alpha(g_2, g_1)}\, U(g_2)U(g_1) \tag{7.11}$$

Cette équation traduit une propriété mathématique : on dit que les opérateurs $U(g)$ forment une *représentation projective* du groupe $\mathcal{G}$. Dans la suite du livre, nous aurons uniquement à considérer deux versions simples de (7.11), l'une où le facteur de phase est $+1$, et dans ce cas, on a affaire à une *représentation vectorielle* de $\mathcal{G}$

$$U(g) = U(g_2)U(g_1) \tag{7.12}$$

et l'autre où le facteur de phase vaut ± 1

$$U(g) = \pm\, U(g_2)U(g_1) \tag{7.13}$$

Nous verrons apparaître ce facteur $\pm$ dans le cas où $\mathcal{G}$ est le groupe des rotations ; les représentations (7.13) de ce groupe sont appelées *représentations spinorielles* du groupe des rotations.

7.2 Générateurs infinitésimaux

7.2.1 Définitions

On distingue deux types de groupes de transformations.

- Les groupes discrets, dont le nombre d'éléments est fini ou dénombrable. Comme cas particuliers simples, on peut citer la parité, ou opération qui change le signe des coordonnés $\vec{r} \to -\vec{r}$ (*cf.* § 7.3.3), ou les groupes cristallographiques qui jouent un rôle important en physique du solide.
- Les groupes continus, dont les éléments sont paramétrés par un ou plusieurs paramètres variant de façon continue[6]. Par exemple, la rotation $\mathcal{R}_z(\theta)$ autour de Oz est paramétrée par l'angle θ qui varie de façon continue entre 0 et 2π.

Les groupes continus intéressants en physique sont les groupes de Lie (exercice 7.5.4), dont un exemple est le groupe des rotations dans un espace à trois dimensions, ou groupe $SO(3)$, le groupe des matrices orthogonales : $\mathcal{R}^T\mathcal{R} = \mathcal{R}\mathcal{R}^T = I$ de déterminant $+1$ dans l'espace à trois dimensions[7] ; A^T désigne l'opérateur transposé de A. Ce groupe va jouer un rôle majeur dans la suite. C'est un groupe à trois paramètres : on peut, par exemple, paramétrer une rotation par deux angles donnant la direction $\hat{n}$ de l'axe de rotation dans un référentiel $Oxyz$ et l'angle de rotation, donc en tout trois angles qui varient de façon continue. Le groupe des rotations possède une infinité de sous-groupes abéliens, les rotations autour d'un axe fixe. Nous allons montrer qu'il suffit de considérer trois sous-groupes abéliens correspondant aux rotations autour de Ox, Oy et Oz : le nombre de ces sous-groupes est égal au nombre de paramètres indépendants. Les rotations de ces sous-groupes sont paramétrées par un angle θ, et selon (7.7), ce paramètre est un paramètre *additif* : le produit de deux rotations d'angles θ_1 et θ_2 est la rotation d'angle $\theta = \theta_1 + \theta_2$. De façon générale, si un groupe de Lie $\mathcal{G}$ est paramétré par n paramètres indépendants, on dira que la dimension du groupe est n, et on pourra se ramener à l'étude de n sous-groupes abéliens (exercice 7.5.4). Soit un sous groupe abélien de $\mathcal{G}$, dont les éléments h sont paramétrés à l'aide d'un paramètre additif α

$$h(\alpha_1 + \alpha_2) = h(\alpha_2)h(\alpha_1) \tag{7.14}$$

6. On peut remarquer que, dans le cas d'un groupe continu, les transformations $U(g)$ doivent nécessairement être unitaires par continuité, si tout élément du groupe peut être relié de façon continue à l'élément neutre e du groupe (en d'autres termes si le groupe est connexe) : en effet $U(e) = I$ est unitaire.

7. La relation $\mathcal{R}^T\mathcal{R} = I$ implique que $\det\mathcal{R} = \pm 1$. Dans la notation $SO(3)$, S indique que l'on doit choisir $\det\mathcal{R} = +1$, O qu'il s'agit du groupe orthogonal et 3 désigne la dimension de l'espace. Si l'on ajoute aux rotations l'opération d'inversion des axes, ou parité, on obtient le groupe $O(3)$, qui inclut aussi les matrices de déterminant -1. Le groupe $SO(3)$ est connexe, mais non $O(3)$: on ne peut pas passer de façon continue de $\det\mathcal{R} = +1$ à $\det\mathcal{R} = -1$.

D'après (7.12), on doit avoir pour les opérateurs $U_h(\alpha)$ qui transforment les vecteurs d'état de $\mathcal{H}$

$$U_h(\alpha_1 + \alpha_2) = U_h(\alpha_2)U_h(\alpha_1) \tag{7.15}$$

Le théorème de Stone (§ 6.3.2) implique qu'il existe alors un opérateur hermitien $T_h = T_h^\dagger$ tel que

$$U_h(\alpha) = \mathrm{e}^{-\mathrm{i}\alpha T_h} \tag{7.16}$$

L'opérateur T_h est appelé *générateur infinitésimal* de la transformation considérée. Comme T_h est hermitien, c'est un bon candidat pour une propriété physique, et de fait, à toutes les transformations dont la liste figure dans l'introduction de ce chapitre correspondent des propriétés physiques fondamentales. En effet, on établit la correspondance suivante entre générateurs infinitésimaux et propriétés physiques pour ces diverses transformations, que nous allons revoir en détail dans la suite de ce chapitre.

- Translations de temps de t : $U(t) = \exp(-\mathrm{i}tH/\hbar)$: $T_h = H =$ hamiltonien : voir le chapitre 4.
- Translations d'espace de $\vec{a} = a\hat{a}$: $U(\vec{a}) = \exp(-\mathrm{i}a(\vec{P}\cdot\hat{a})/\hbar)$: $T_h = \vec{P}\cdot\hat{a} =$ composante suivant $\hat{a}$ de l'impulsion $\vec{P}$.
- Rotations autour d'un axe $\hat{n}$: $U_{\hat{n}}(\theta) = \exp(-\mathrm{i}\theta(\vec{J}\cdot\hat{n})/\hbar)$: $T_h = \vec{J}\cdot\hat{n} =$ composante suivant $\hat{n}$ du moment angulaire $\vec{J}$.
- Transformation de Galilée de vitesse $\vec{v}$: $U(\vec{v}) = \exp(-\mathrm{i}(\vec{v}\cdot\vec{G})/\hbar)$: $\vec{G} = -m\vec{R}$, $\vec{R} =$ position, m étant la masse.

Dans chaque cas, la présence de $\hbar$ dans l'exponentielle assure que l'exposant est une quantité sans dimensions. Si l'on choisit précisément $\hbar$, et non $\hbar$ que multiplie une constante numérique, alors les expressions précédentes *définissent* les opérateurs représentant les propriétés physiques énergie, impulsion, moment angulaire et position. En fait, ces expressions donnent la définition *la plus générale* de ces opérateurs.

7.2.2 Lois de conservation

Nous allons montrer qu'aux lois de conservation de la physique classique, en présence d'une symétrie, correspondent en physique quantique des lois de conservation pour les valeurs moyennes de grandeurs physiques. Généralisons d'abord (4.26) au cas où l'opérateur A dépend explicitement du temps. Au membre de droite de (4.26), on doit ajouter

$$\langle\varphi(t)\Big|\frac{\partial A}{\partial t}\Big|\varphi(t)\rangle = \Big\langle\frac{\partial A}{\partial t}\Big\rangle_\varphi$$

et cette équation donne la forme générale du *théorème d'Ehrenfest*

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_\varphi(t) = \frac{\mathrm{i}}{\hbar}\langle[H,A]\rangle_\varphi + \Big\langle\frac{\partial A}{\partial t}\Big\rangle_\varphi} \tag{7.17}$$

Lorsque l'opérateur A est indépendant du temps, $(\partial A/\partial t) = 0$ et l'on retrouve (4.26)

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_\varphi(t) = \frac{\mathrm{i}}{\hbar}\langle[H,A]\rangle_\varphi \tag{7.18}$$

Comme cette égalité est valable quel que soit $|\varphi\rangle$, nous obtenons le théorème suivant (nous supposons H indépendant du temps).
Théorème de conservation de la valeur moyenne. Lorsque la propriété physique A est indépendante du temps, la condition $\mathrm{d}\langle A\rangle/\mathrm{d}t = 0$ implique $[H,A] = 0$ et réciproquement.

$$\text{Si } \frac{\partial A}{\partial t} = 0, \quad \frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_\varphi = 0 \Longleftrightarrow [H,A] = 0 \ \blacksquare \tag{7.19}$$

Comme application, supposons que les propriétés d'un système physique soient invariantes par toute translation d'espace. Ce sera le cas, par exemple, pour un système isolé de deux particules dont l'énergie potentielle dépend uniquement de la différence de leurs positions $(\vec{r}_1 - \vec{r}_2)$. La valeur moyenne du hamiltonien doit être la même dans l'état $|\varphi\rangle$ et l'état $|\varphi_{\vec{a}}\rangle = \exp[-\mathrm{i}(\vec{P}\cdot\vec{a})/\hbar]|\varphi\rangle$ obtenu par translation de $\vec{a}$, où $\vec{a}$ est un vecteur arbitraire

$$\langle\varphi_{\vec{a}}|H|\varphi_{\vec{a}}\rangle = \langle\varphi|\exp\left(\mathrm{i}\,\frac{\vec{P}\cdot\vec{a}}{\hbar}\right) H \exp\left(-\,\mathrm{i}\,\frac{\vec{P}\cdot\vec{a}}{\hbar}\right)|\varphi\rangle = \langle\varphi|H|\varphi\rangle$$

Faisant tendre $\vec{a}$ vers zéro, on en déduit

$$\boxed{\text{Invariance par translation d'espace} \Longleftrightarrow [H,\vec{P}] = 0} \tag{7.20}$$

La notation $[H,\vec{P}] = 0$ indique que les trois composantes de l'impulsion commutent avec H. D'après (7.18), cette équation implique que la valeur moyenne $\langle\vec{P}\rangle$ de $\vec{P}$ est indépendante du temps : l'invariance par translation entraîne la conservation de l'impulsion (en valeur moyenne). Un raisonnement identique montre que

$$\boxed{\text{Invariance par rotation} \Longleftrightarrow [H,\vec{J}] = 0} \tag{7.21}$$

La valeur moyenne $\langle\vec{J}\rangle$ de $\vec{J}$ est indépendante du temps : l'invariance par rotation entraîne la conservation du moment angulaire (en valeur moyenne).

Il est également utile de faire les deux remarques suivantes.

- Si $[H,A] = 0$, A et H peuvent être diagonalisés simultanément, et en particulier, on peut choisir les états stationnaires parmi les vecteurs propres de A.
- La condition $[H,A] = 0$ implique que A commute avec l'opérateur d'évolution $U(t-t_0)$ (4.20). Si $|\varphi(t_0)\rangle$ est vecteur propre de A au temps t_0

$$A|\varphi(t_0)\rangle = a|\varphi(t_0)\rangle$$

alors $|\varphi(t)\rangle$ est vecteur propre de A avec la même valeur propre

$$A|\varphi(t)\rangle = AU(t-t_0)|\varphi(t_0)\rangle = U(t-t_0)A|\varphi(t_0)\rangle = a|\varphi(t)\rangle$$

La valeur propre a est conservée : c'est une constante du mouvement. On aurait pu déduire ce résultat directement de (7.19), puisque dans ce cas $\langle A\rangle = a$.

7.2.3 Relations de commutation des générateurs infinitésimaux

On peut déterminer la plupart des propriétés d'un groupe de Lie en examinant le voisinage de l'identité, plus précisément en étudiant les relations de commutation de ses générateurs infinitésimaux : l'ensemble de ces relations de commutation constitue *l'algèbre de Lie* du groupe (exercice 7.5.4). Cependant, deux groupes de Lie isomorphes dans le voisinage de l'identité peuvent différer par des propriétés topologiques globales : nous en verrons bientôt un exemple. Examinons plus en détail le cas du groupe des rotations[8]. L'opérateur de rotation $\mathcal{R}_{\hat{n}}(\theta)$ d'un angle θ autour d'un axe $\hat{n}$ est un opérateur orthogonal de l'espace à trois dimensions : $\mathcal{R}^T\mathcal{R} = \mathcal{R}\,\mathcal{R}^T = I$. Les rotations $\mathcal{R}_{\hat{n}}(\theta)$ forment un sous groupe abélien du groupe des rotations, et toujours d'après le théorème de Stone, on peut écrire

$$\mathcal{R}_{\hat{n}}(\theta) = \exp\Big(-\mathrm{i}\theta(\vec{T}\cdot\hat{n})\Big) \tag{7.22}$$

où $\vec{T}\cdot\hat{n}$ est un opérateur hermitien : $\mathcal{R}$ étant orthogonal et réel, est aussi unitaire. Dans une telle rotation, un vecteur $\vec{V}$ se transforme en $\vec{V}'$ (figure 7.3)

$$\vec{V}' = (1-\cos\theta)(\hat{n}\cdot\vec{V})\hat{n} + \cos\theta\,\vec{V} + \sin\theta(\hat{n}\times\vec{V}) \tag{7.23}$$

On peut écrire cette loi de tranformation sous forme matricielle

$$V_i' = \sum_{j=1}^{3}[\mathcal{R}_{\hat{n}}(\theta)]_{ij}V_j \tag{7.24}$$

La détermination explicite de la matrice $[\mathcal{R}_{\hat{n}}(\theta)]_{ij}$ est proposée dans l'exercice 7.5.1. Nous n'aurons pas à nous en servir, car nous allons prendre la limite $\theta \to 0$, c'est-à-dire la limite des rotations infinitésimales

$$\vec{V}' = \vec{V} + \theta(\hat{n}\times\vec{V}) + O(\theta)^2 \tag{7.25}$$

Le développement de l'exponentielle dans (7.22) et la comparaison avec (7.25) donnent

$$(\vec{T}\cdot\hat{n})\vec{V} = \mathrm{i}\begin{pmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{pmatrix}\begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix}$$

8. Sauf mention explicite du contraire, il s'agira toujours du groupe $SO(3)$ des rotations dans l'espace euclidien à trois dimensions.

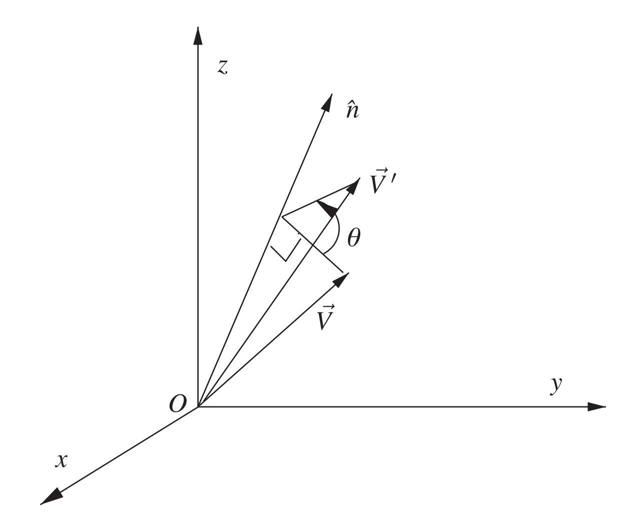


FIG. 7.3 – Rotation de θ d'un vecteur $\vec{V}$ autour d'un axe $\hat{n}$.

et par identification les opérateurs hermitiens T_x, T_y et T_z

$$T_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix} \quad T_y = \begin{pmatrix} 0 & 0 & \mathrm{i} \\ 0 & 0 & 0 \\ -\mathrm{i} & 0 & 0 \end{pmatrix} \quad T_z = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{7.26}$$

Lorsque θ est fini, on peut aisément calculer l'exponentielle dans (7.22) en remarquant que $(\vec{T}\cdot\hat{n})^3 = \vec{T}\cdot\hat{n}$ (exercice 7.5.1) et vérifier que l'on retrouve bien (7.23). Un calcul direct (exercice 7.5.1) permet de montrer les relations de commutation[9] suivantes, qui constituent l'algèbre de Lie de $SO(3)$

$$[T_x, T_y] = \mathrm{i}\,T_z \qquad [T_y, T_z] = \mathrm{i}\,T_x \qquad [T_z, T_x] = \mathrm{i}\,T_y \tag{7.27}$$

ou en reprenant les notations de (3.52)

$$\boxed{[T_i, T_j] = \mathrm{i}\sum_k \varepsilon_{ijk} T_k} \tag{7.28}$$

Nous allons donner une démonstration plus rapide et surtout plus instructive de (7.27) en utilisant l'expression suivante pour une rotation d'angle θ autour d'un axe $\hat{n}(\phi)$ du plan yOz, obtenu à partir de l'axe Oy par une rotation d'angle ϕ autour de Ox (figure 7.4)

$$\mathcal{R}_{\hat{n}(\phi)}(\theta) = \mathcal{R}_x(\phi)\mathcal{R}_y(\theta)\mathcal{R}_x(-\phi) \tag{7.29}$$

9. Il suffit, bien sûr, de montrer la première relation, les deux autres s'en déduisant par permutation circulaire.

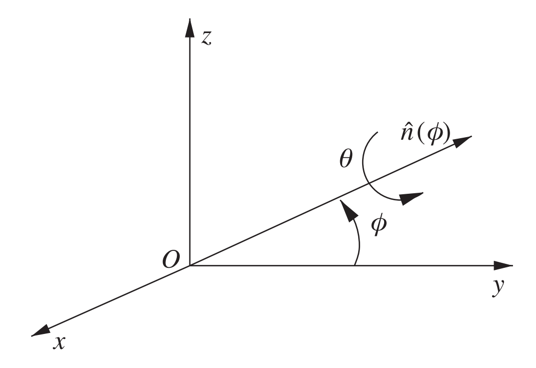


FIG. 7.4 – Rotation $\mathcal{R}_{\hat{n}(\phi)}(\theta)$.

En effet, la rotation $\mathcal{R}_x(-\phi)$ amène d'abord l'axe $\hat{n}(\phi)$ sur Oy ; on effectue ensuite la rotation d'angle θ autour de Oy et on revient enfin à la position initiale de l'axe par la rotation $\mathcal{R}_x(\phi)$. Exprimons $\mathcal{R}_{\hat{n}(\phi)}(\theta)$ et $\mathcal{R}_y(\theta)$ sous forme exponentielle (7.22) et développons au premier ordre en θ

$$\vec{T} \cdot \hat{n}(\phi) = \cos\phi\, T_y + \sin\phi\, T_z = \mathrm{e}^{-\mathrm{i}\phi T_x}\, T_y \mathrm{e}^{\,\mathrm{i}\phi T_x}$$

En développant au premier ordre en ϕ, on obtient

$$[T_x, T_y] = \mathrm{i}\, T_z$$

et les deux autres relations de commutation (7.27) s'en déduisent par permutation circulaire.

Considérons maintenant les opérateurs qui effectuent des rotations sur les états physiques dans $\mathcal{H}$. Nous avons vu que l'opérateur qui effectue une rotation d'angle θ autour d'un axe $\hat{n}$ est

$$\boxed{U_{\hat{n}}(\theta) = \exp\Big(-\mathrm{i}\theta\, \frac{\vec{J} \cdot \hat{n}}{\hbar}\Big)} \tag{7.30}$$

où $\vec{J} \cdot \hat{n}$ est le moment angulaire suivant $\hat{n}$. Comme ces opérateurs forment une représentation du groupe des rotations, on déduit de (7.12) et de (7.29)

$$U_{\hat{n}(\phi)}(\theta) = U_x(\phi) U_y(\theta) U_x(-\phi)$$

En développant comme précédemment les exponentielles au premier ordre en θ puis en ϕ on obtient les relations de commutation du moment angulaire

$$[J_x, J_y] = \mathrm{i}\hbar J_z \qquad [J_y, J_z] = \mathrm{i}\hbar J_x \qquad [J_z, J_x] = \mathrm{i}\hbar J_y \tag{7.31}$$

ou

$$\boxed{[J_i, J_j] = \mathrm{i}\hbar \sum_k \varepsilon_{ijk}\, J_k} \tag{7.32}$$

Les relations de commutation des J_i sont donc, au facteur $\hbar$ près, identiques à celles des T_i. *Les générateurs infinitésimaux des rotations dans $\mathcal{H}$ ont des relations de commutation identiques à celles des générateurs infinitésimaux du groupe des rotations dans l'espace ordinaire.* Notre démonstration des relations (7.31) ou (7.32) souligne leur origine géométrique.

Les relations de commutation des opérateurs scalaires et vectoriels avec $\vec{J}$ sont d'une grande importance pratique. Un *opérateur scalaire* $\mathcal{S}$ est un opérateur dont la valeur moyenne est invariante dans une rotation. Si $U(\mathcal{R})$ est l'opérateur qui effectue la rotation $\mathcal{R}$ dans l'espace des états

$$|\varphi_{\mathcal{R}}\rangle = U(\mathcal{R})|\varphi\rangle$$

nous devons avoir

$$\langle\varphi_{\mathcal{R}}|\mathcal{S}|\varphi_{\mathcal{R}}\rangle = \langle\varphi|U^\dagger(\mathcal{R})\mathcal{S}U(\mathcal{R})|\varphi\rangle = \langle\varphi|\mathcal{S}|\varphi\rangle$$

et par conséquent pour une rotation $R_{\hat{n}}(\theta)$

$$\exp\left(\mathrm{i}\,\theta\frac{\vec{J}\cdot\hat{n}}{\hbar}\right)\mathcal{S}\exp\left(-\,\mathrm{i}\theta\,\frac{\vec{J}\cdot\hat{n}}{\hbar}\right) = \mathcal{S}$$

Prenant θ infinitésimal, nous constatons que $\mathcal{S}$ commute avec $\vec{J}$

$$\boxed{[\vec{J},\mathcal{S}] = 0} \tag{7.33}$$

Un opérateur scalaire commute avec le moment angulaire.

Un raisonnement analogue permet d'établir les relations de commutation de $\vec{J}$ avec $\vec{R}$ ou $\vec{P}$, et plus généralement avec tous les *opérateurs vectoriels*. Par définition, un opérateur vectoriel $\vec{V}$ est un opérateur dont la valeur moyenne se transforme par rotation suivant la loi (7.24). Nous devons donc avoir

$$\langle\varphi_{\mathcal{R}}|V_i|\varphi_{\mathcal{R}}\rangle = \langle\varphi|U^\dagger(\mathcal{R})V_iU(\mathcal{R})|\varphi\rangle = \sum_{j=1}^{3}\mathcal{R}_{ij}\,\langle\varphi|V_j|\varphi\rangle$$

et par conséquent pour une rotation $\mathcal{R}_{\hat{n}}(\theta)$

$$\exp\left(\mathrm{i}\theta\,\frac{\vec{J}\cdot\hat{n}}{\hbar}\right)V_i\exp\left(-\,\mathrm{i}\theta\,\frac{\vec{J}\cdot\hat{n}}{\hbar}\right) = \sum_{j=1}^{3}[\mathcal{R}_{\hat{n}}(\theta)]_{ij}V_j \tag{7.34}$$

Prenons $\hat{n} = \hat{x}$ et θ infinitésimal ; d'après (7.25), $\vec{V}'$ a pour composantes

$$(V_x,\ V_y - \theta V_z,\ V_z + \theta V_y)$$

et nous avons donc, par exemple pour la composante $i = y$ de (7.34)

$$\left(I + \frac{\mathrm{i}}{\hbar}\,\theta J_x\right)V_y\left(I - \frac{\mathrm{i}}{\hbar}\,\theta J_x\right) = V_y - \theta V_z$$

soit $i[J_x, V_y] = -\hbar V_z$ et en examinant les autres composantes

$$[J_x, V_x] = 0 \qquad [J_x, V_y] = \mathrm{i}\hbar V_z \qquad [J_x, V_z] = -\mathrm{i}\hbar V_y$$

ou sous forme générale

$$\boxed{[J_i, V_j] = \mathrm{i}\hbar \sum_k \varepsilon_{ijk} V_k} \tag{7.35}$$

Ces relations sont valables en particulier pour l'opérateur position $\vec{R}$ et l'opérateur impulsion $\vec{P}$, qui sont des opérateurs vectoriels.

Le lecteur attentif aura remarqué que les relations de commutation (3.53) du spin 1/2, $\vec{S} = \frac{1}{2}\hbar\vec{\sigma}$, sont identiques à (7.31) ou (7.32), et que le spin 1/2 est donc un moment angulaire. Donnons quelques indications permettant de comprendre cette identification, sans toutefois entrer dans des détails mathématiques qui nous entraîneraient trop loin. L'algèbre de Lie (3.52) des matrices de Pauli est celle du groupe $SU(2)$ des matrices 2×2 unitaires et de déterminant $+1$ (exercice 7.5.2). L'algèbre de Lie de $SU(2)$ et celle de $SO(3)$ sont identiques : les deux groupes coïncident dans le voisinage de l'identité. Cependant, les deux groupes ne sont pas *globalement* identiques : on le voit en considérant une rotation de 2π autour d'un axe $\hat{n}$. Compte tenu de (3.59),

$$\exp\Big(-\mathrm{i}\frac{\theta}{2}\vec{\sigma}\cdot\hat{n}\Big) = -I \quad \text{pour} \quad \theta = 2\pi$$

On retrouve l'identité seulement pour $\theta = 4\pi$! À la rotation identité de $SO(3)$ correspondent donc *deux* éléments de $SU(2)$, $+I$ et $-I$. La correspondance entre $SU(2)$ et $SO(3)$ est un homomorphisme, qui à deux éléments de $SU(2)$ fait correspondre un élément de $SO(3)$, et on a donc pour un spin 1/2 une représentation projective (7.13) du groupe des rotations. Cette propriété découle de ce que le groupe $SO(3)$ est connexe mais non simplement connexe[10] : une courbe continue fermée dans l'espace des paramètres du groupe ne peut pas toujours être déformée continûment en un point. Cette propriété est visible dans les rotations de l'espace ordinaire[11] ; elle n'est pas propre à la mécanique

10. Un disque dans le plan est simplement connexe. Perçons un trou dans ce disque : alors la région du plan ainsi obtenue n'est plus simplement connexe, car une courbe encerclant le trou ne peut plus être déformée en un point.

11. *cf.* Lévy-Leblond et Balibar [1984], chapitre 3.D ; l'argument est dû à Dirac.

quantique comme on a parfois tendance à l'affirmer[12] : la véritable rotation identité pour un objet *en relation avec son environnement* n'est pas la rotation de 2π mais la rotation de 4π !

7.3 Relations de commutation canoniques

7.3.1 Cas de la dimension $d = 1$

Plaçons-nous d'abord à une dimension, sur l'axe des x, et soit X l'opérateur position. Considérons une particule dans un état $|\varphi\rangle$ où cette particule est localisée au voisinage d'une position moyenne x_0, avec une dispersion Δx

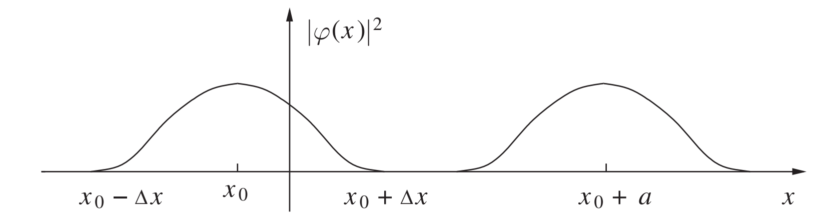


FIG. 7.5 – Particules localisées au voisinage de $x = x_0$ et $x = x_0 + a$.

$$\langle\varphi|X|\varphi\rangle = \langle X\rangle = x_0 \qquad \langle\varphi|(X - x_0)^2|\varphi\rangle = (\Delta x)^2 \tag{7.36}$$

Elle est, par exemple, localisée dans l'intervalle $[x_0 - \Delta x, x_0 + \Delta x]$ (figure 7.5). Si nous appliquons à cet état une translation a

$$|\varphi\rangle \to |\varphi_a\rangle = \exp\Big(-\mathrm{i}\,\frac{Pa}{\hbar}\Big)|\varphi\rangle = U(a)|\varphi\rangle$$

12. Un mot sur les conditions où les représentations projectives sont nécessaires. Deux cas peuvent se présenter. (i) Comme pour la correspondance entre $SU(2)$ et $SO(3)$, la nécessité d'une représentation projective vient de propriétés topologiques globales. Le facteur de phase dans (7.11) prend alors des valeurs discrètes, comme dans (7.13). (ii) Si

$$[T_i, T_j] = \mathrm{i}\sum_k C_{ijk}T_k$$

est l'algèbre de Lie du groupe (dont (7.28) pour $SO(3)$ est un exemple, voir aussi l'exercice 7.5.4), il se peut que l'on puisse construire une autre algèbre de Lie dont le second membre diffère par un multiple de l'identité :

$$[T'_i, T'_j] = \mathrm{i}\sum_k C_{ijk}T'_k + \mathrm{i}D_{ij}I \qquad D_{ij} = -D_{ji}$$

C'est ce que l'on appelle une extension centrale de l'algèbre de Lie initiale. Si le terme $D_{ij}I$ peut être éliminé par une redéfinition des générateurs infinitésimaux T'_i, alors il n'existe que des représentations vectorielles (avec éventuellement des facteurs de phase discrets dus aux propriétés topologiques globales comme dans (i)). Dans le cas contraire, par exemple celui du groupe de Galilée (exercice 7.5.7), il existe des représentations projectives où le facteur de phase varie de façon continue : *cf.* par exemple Weinberg [1995], chapitre 2.

où P est l'opérateur impulsion et $U(a)$ l'opérateur de translation

$$\boxed{U(a) = \exp\Big(-\,\mathrm{i}\,\frac{Pa}{\hbar}\Big) \qquad U^{-1}(a) = U^{\dagger}(a) = \exp\Big(\mathrm{i}\,\frac{Pa}{\hbar}\Big)} \tag{7.37}$$

alors la particule sera après translation localisée dans l'intervalle $[x_0 + a - \Delta x, x_0 + a + \Delta x]$

$$\langle X \rangle_a = \langle \varphi_a | X | \varphi_a \rangle = \langle \varphi | U^{-1}(a) X U(a) | \varphi \rangle = x_0 + a = \langle X \rangle + a$$

Comme l'état $|\varphi\rangle$ est arbitraire, l'égalité des valeurs moyennes entraîne celle des opérateurs

$$U^{-1}(a) X U(a) = X + aI \tag{7.38}$$

et en faisant tendre a vers zéro, nous obtenons la *relation de commutation canonique* (RCC) entre X et P

$$\boxed{[X, P] = \mathrm{i}\hbar I} \tag{7.39}$$

Comme application, calculons le commutateur entre P et une fonction quelconque $f(X)$. Développons $f(X)$ en série de Taylor

$$f(X) = c_0 + c_1 X^2 + \ldots + c_n X^n + \ldots$$

D'après (7.38),

$$U^{-1}(a) X^2 U(a) = U^{-1}(a) X U(a) U^{-1}(a) X U(a) = (X + aI)^2$$

et ceci se généralise immédiatement à X^n

$$U^{-1}(a) X^n U(a) = (X + aI)^n$$

Nous obtenons donc

$$\boxed{U^{-1}(a) f(X) U(a) = f(X + aI)} \tag{7.40}$$

Selon une technique maintenant éprouvée, nous faisons tendre a vers zéro

$$[P, f(X)] = -\mathrm{i}\hbar \frac{\partial f(X)}{\partial X} \tag{7.41}$$

Comme cas particulier de (7.40), nous pouvons choisir $f(X) = \exp(\mathrm{i}\beta X)$, β réel, et nous obtenons les relations de commutation canoniques sous la forme de Weyl

$$\exp\Big(\mathrm{i}\,\frac{Pa}{\hbar}\Big) \exp(\mathrm{i}\beta X) \exp\Big(-\,\mathrm{i}\,\frac{Pa}{\hbar}\Big) = \exp(\mathrm{i}\beta X) \exp(\mathrm{i}\beta a) \tag{7.42}$$

La forme de Weyl est plus intéressante mathématiquement que (7.39), car les opérateurs unitaires qui interviennent dans (7.42) sont bornés (§ 6.2.1), contrairement aux opérateurs X et P.

On déduit immédiatement de (7.39) l'inégalité de Heisenberg sur les dispersions en position et impulsion ; en effet, d'après (4.10)

$$\Delta x \, \Delta p = \sqrt{\langle (X - x)^2 \rangle} \sqrt{\langle (P - p)^2 \rangle} \geq \frac{1}{2}\,\hbar \tag{7.43}$$

7.3.2 Réalisation explicite et commentaires

Une réalisation explicite, ou *représentation* des relations de commutation canoniques (7.39) peut être donnée dans l'espace $L^{(2)}(\mathbb{R})$ des fonctions $\varphi(x)$ différentiables de carré sommable sur la droite dans l'intervalle $]-\infty,+\infty[$. Cette représentation est

$$\boxed{(X\varphi)(x) = x\varphi(x) \qquad (P\varphi)(x) = -\mathrm{i}\hbar\frac{\partial\varphi}{\partial x}} \tag{7.44}$$

Dans ces équations, $(X\varphi)$ et $(P\varphi)$ sont des symboles de fonctions, par exemple $(X\varphi)(x) = g(x)$ et $(P\varphi)(x) = h(x)$. Vérifions (7.44)

$$([XP - PX]\varphi)(x) = -\mathrm{i}\hbar x\frac{\partial\varphi}{\partial x} + \mathrm{i}\hbar\frac{\partial}{\partial x}(x\varphi(x) = \mathrm{i}\hbar\varphi(x)$$

ou

$$([X, P]\varphi)(x) = \mathrm{i}\hbar\varphi(x)$$

Il est légitime de se poser la question de l'unicité de la représentation (7.44) des relations de commutation canoniques : l'équation (7.44) est-elle une solution unique de (7.39) ? Évidemment, deux représentations ne doivent pas être considérées comme distinctes si elles sont reliées par une transformation unitaire, qui est un simple changement de base orthonormée dans $\mathcal{H}$. Soit U un opérateur unitaire. Les opérateurs P' et X' obtenus par transformation unitaire

$$P' = U^\dagger PU \qquad X' = U^\dagger XU$$

obéissent aussi aux relations de commutation canoniques

$$[X', P'] = U^\dagger XUU^\dagger PU - U^\dagger PUU^\dagger XU = U^\dagger[X, P]U = \mathrm{i}\hbar I$$

La représentation (X', P') des relations de commutation canoniques est dite unitairement équivalente à la représentation (X, P). La grande importance de (7.44) vient du théorème suivant, que nous énonçons sans démonstration. *Théorème d'équivalence unitaire de von Neumann.* Toutes les représentations des relations de commutation canoniques sous la forme de Weyl[13] (7.42) sont *unitairement équivalentes* à la représentation (7.44) sur $L^{(2)}(\mathbb{R})$. De plus, cette représentation est irréductible, c'est-à-dire que tout opérateur sur $\mathcal{H}$ peut s'écrire comme une fonction de X et de P. Tout opérateur commutant avec X (resp. P) est une fonction de X (resp. P). Tout opérateur commutant avec X et P est multiple de l'identité I.■

Ce théorème implique que nous n'avons pas à nous préoccuper du choix de la représentation de (7.39), puisque deux choix différents seront reliés par une transformation unitaire.

13. Cette précision est importante : sinon les opérateurs A_C du § 6.2.2 permettraient de construire un contre-exemple au théorème !

À trois dimensions, les opérateurs position $\vec{R}$ et impulsion $\vec{P}$ sont des opérateurs vectoriels de composantes X, Y, Z et P_x, P_y, P_z, que l'on note collectivement X_i et P_i, $i = x, y, z$. Les composantes différentes de $\vec{R}$ et $\vec{P}$ commutent, et seules les composantes identiques ont des relations de commutation non nulles

$$\boxed{[X_i, P_j] = \mathrm{i}\hbar\, \delta_{ij} I} \tag{7.45}$$

7.3.3 L'opération parité

L'opération parité consiste à inverser le signe des coordonnées : $\vec{r} \to -\vec{r}$. On peut aussi la voir comme la combinaison d'une réflexion par rapport à un plan suivie d'une rotation de π autour d'un axe perpendiculaire à ce plan. Choisissons par exemple le plan xOy et appelons M la réflexion par rapport à ce plan, $\mathcal{R}_z(\pi)$ la rotation autour de Oz

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \underset{M}{\longrightarrow} \begin{pmatrix} x \\ y \\ -z \end{pmatrix} \underset{\mathcal{R}_z(\pi)}{\longrightarrow} \begin{pmatrix} -x \\ -y \\ -z \end{pmatrix} \tag{7.46}$$

Comme l'invariance par rotation est en général valable, on exprime de façon imagée l'invariance par parité en disant que *l'image dans un miroir d'une expérience de physique doit apparaître comme physiquement possible*. L'opération parité agit différemment sur les vecteurs proprement dits, ou *vecteurs polaires* comme la position $\vec{r}$, l'impulsion $\vec{p}$ ou le champ électrique $\vec{E}$

$$\vec{r} \to -\vec{r} \qquad \vec{p} \to -\vec{p} \qquad \vec{E} \to -\vec{E} \tag{7.47}$$

et sur les *pseudovecteurs*, ou *vecteurs axiaux*, comme le moment angulaire $\vec{\jmath}$ ou le champ magnétique $\vec{B}$, qui sont associés à un sens de rotation autour d'un axe, et non à une direction

$$\vec{\jmath} \to \vec{\jmath} \qquad \vec{B} \to \vec{B} \tag{7.48}$$

Rappelons que le produit vectoriel de deux vecteurs polaires est un vecteur axial[14] : $\vec{\jmath} = \vec{r} \times \vec{p}$ est un vecteur axial.

Les interactions faibles (*cf.* § 1.1.4) ne respectent pas l'invariance par parité : ceci a été montré pour la première fois par C.S. Wu, en utilisant la désintégration β (1.4) de noyaux de ^{60}Co polarisés en un état excité du ^{60}Ni

$$^{60}\mathrm{Co} \to {}^{60}\mathrm{Ni}^* + e^- + \overline{\nu}$$

La valeur moyenne du moment angulaire $\langle \vec{J} \rangle$ du ^{60}Co a une orientation fixée (figure 7.6). On constate que les électrons de la désintégration sont émis de

14. L'existence de vecteurs axiaux est une particularité de l'espace à trois dimensions, $d = 3$. Un vecteur axial est en fait un tenseur antisymétrique de rang 2, qui a $d(d-1)/2$ composantes en général. Pour $d = 3$, ce nombre de composantes est 3, ce qui permet de lui faire correspondre un (pseudo)vecteur. À 4 dimensions, une telle identification n'est plus possible : un tenseur antisymétrique de rang 2 comme le champ électromagnétique a 6 composantes.

façon préférentielle dans la direction opposée à celle du moment angulaire : si $\vec{P}$ est l'impulsion des électrons, $\langle \vec{J} \cdot \vec{P} \rangle < 0$. Mais $\langle \vec{J} \cdot \vec{P} \rangle$, valeur moyenne du produit scalaire d'un vecteur polaire et d'un vecteur axial, est un pseudoscalaire, qui change de signe dans une opération parité. L'image de l'expérience dans un miroir (figure 7.6) n'apparaît pas comme physiquement possible : dans le miroir les sens de rotation sont inversés, et les électrons partent préférentiellement dans la direction de $\vec{J}$.

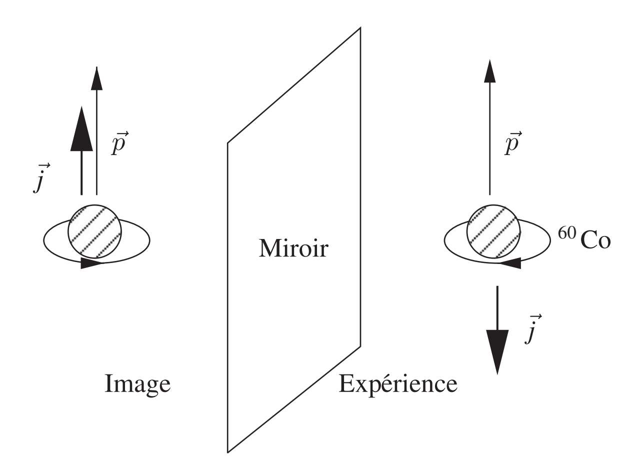


Fig. 7.6 – L'expérience de désintégration du cobalt polarisé et son image dans un miroir.

Le groupe $\mathcal{G}$ correspondant à l'opération parité est le groupe multiplicatif à deux élements $\{+1, -1\}$, ou groupe Z_2. Comme on ne peut pas relier continûment -1 à l'identité, il nous faut trouver un argument pour décider si l'opérateur Π, qui représente l'opération parité dans l'espace des états, est unitaire ou antiunitaire. Soit χ et φ deux vecteurs arbitraires et (χ, φ) leur produit scalaire (nous revenons provisoirement aux notations des mathématiciens). Si la parité est une symétrie,

$$|(\Pi\chi, \Pi\varphi)| = |(\chi, \varphi)|$$

Comme dans l'opération parité, les opérateurs position et impulsion doivent se transformer tous deux comme des vecteurs :

$$\vec{R} \to \Pi^{-1}\vec{R}\,\Pi = -\vec{R} \qquad \vec{P} \to \Pi^{-1}\vec{P}\,\Pi = -\vec{P} \tag{7.49}$$

leur commutateur est inchangé

$$\Pi[X_i, P_j]\Pi^{-1} = \mathrm{i}\hbar\delta_{ij}I$$

Examinons l'élément de matrice

$$\begin{aligned}(\Pi\chi, \Pi[X_i, P_j]\varphi) &= (\Pi\chi, \Pi[X_i, P_j]\Pi^{-1}\Pi\varphi) \\ &= (\Pi\chi, \mathrm{i}\hbar\delta_{ij}\Pi\varphi) = \mathrm{i}\hbar\delta_{ij}(\Pi\chi, \Pi\varphi)\end{aligned} \tag{7.50}$$

Mais, on a également

$$\begin{aligned}(\Pi\chi, \Pi[X_i, P_j]\varphi) &= (\Pi\chi, \Pi\, \mathrm{i}\hbar\delta_{ij}\varphi) \\ &= \mathrm{i}\hbar\delta_{ij}(\Pi\chi, \Pi\varphi)\end{aligned} \tag{7.51}$$

si l'on suppose que Π est unitaire. En effet, pour un opérateur unitaire

$$(U\chi, U\mathrm{i}\varphi) = (\chi, \mathrm{i}\varphi) = \mathrm{i}(\chi, \varphi)$$

tandis que pour un opérateur antiunitaire

$$(U\chi, U\mathrm{i}\varphi) = (\mathrm{i}\varphi, \chi) = -\mathrm{i}(\varphi, \chi)$$

Les équations (7.50) et (7.51) sont compatibles uniquement si Π est unitaire. En revanche, si au lieu de la parité Π on considère le renversement du sens du temps Θ : $\vec{R} \to \vec{R}$ et $\vec{P} \to -\vec{P}$ (*cf.* annexe A.2), alors

$$\Theta[X_i, P_j]\Theta^{-1} = -[X_i, P_j] = -\mathrm{i}\hbar\delta_{ij}$$

et ce changement de signe entraîne que Θ est antiunitaire.

Si la parité est une symétrie, ce qui dans l'état actuel de nos connaissances est le cas pour les interactions fortes et les interactions électromagnétiques, alors Π doit commuter avec le hamiltonien : $[\Pi, H] = 0$. Comme $\Pi^2 = I$, puisque deux opérations parité successives ramènent le système d'axes à sa position initiale, les valeurs propres de Π sont ± 1. Comme Π et H commutent, on peut trouver un système de vecteurs propres communs $|\varphi_\pm\rangle$ à H et à Π

$$H|\varphi_\pm\rangle = E_\pm|\varphi_\pm\rangle \qquad \Pi|\varphi_\pm\rangle = \pm|\varphi_\pm\rangle \tag{7.52}$$

Les états $|\varphi_+\rangle$ sont dits de *parité paire* et les états $|\varphi_-\rangle$ de *parité impaire.*

7.4 Invariance galiléenne

7.4.1 Hamiltonien en dimension $d = 1$

Nous allons maintenant examiner les conséquences de la dernière invariance que nous n'avons pas encore exploitée, l'invariance par changement de référentiel d'inertie. Nous nous limitons d'abord à une dimension, une particule sur l'axe des x. Les équations de la physique non relativiste doivent garder la même forme dans la transformation de Galilée

$$x' = x - vt \tag{7.53}$$

qui fait passer d'un référentiel d'inertie à un autre référentiel d'inertie se déplaçant à la vitesse v par rapport au premier. La loi de transformation (7.53) correspond au point de vue passif du changement d'axes. Afin de garder la cohérence avec les sections précédentes, nous allons choisir le point de vue actif, qui consiste à modifier la vitesse de toutes les particules de v ; en bon français[15], on "booste" toutes les particules de v. Si la position, la vitesse,

15. Il n'existe pas de bonne traduction de ce terme dont l'origine est le "booster" d'une fusée, voir aussi la section 19.1.

l'impulsion p et l'énergie cinétique K initiales d'une particule classique de masse m sont

$$x, \quad \dot{x}, \quad p = m\dot{x}, \quad K = \frac{1}{2}\, m\dot{x}^2$$

ces mêmes variables deviendront après le "boost" v

$$x' = x + vt, \qquad \dot{x}' = \dot{x} + v, \qquad p' = m\dot{x}', \qquad K' = \frac{1}{2}\, m(\dot{x}')^2 \tag{7.54}$$

Contrairement au cas des translations et des rotations, l'énergie n'est pas invariante dans une transformation de Galilée. On doit seulement exiger que la forme des équations de la physique reste invariante, ce que l'on appelle aussi la *covariance* de ces équations par rapport à cette transformation.

Venons-en maintenant au cas quantique, en nous plaçant au temps $t = 0$, ce qui correspond à une transformation de Galilée instantanée. La loi de transformation des vecteurs d'état dans une transformation de Galilée sera une transformation unitaire $U(v)$

$$U(v) = \exp\Big(-\mathrm{i}\,\frac{vG}{\hbar}\Big) \tag{7.55}$$

où $G = G^\dagger$ est le générateur infinitésimal des transformations de Galilée. Les transformations de Galilée à une dimension forment un groupe additif, puisque la composition de deux transformations de vitesses v et v' est une transformation de vitesse $v'' = v + v'$. Une fois de plus, le théorème de Stone nous garantit l'existence d'un générateur infinitésimal hermitien G. Si $\langle A\rangle$ est la valeur moyenne d'une propriété physique dans l'état $|\varphi\rangle$, sa valeur moyenne $\langle A\rangle_v$ dans l'état transformé $|\varphi_v\rangle = U(v)|\varphi\rangle$ sera

$$\langle\varphi_v|A|\varphi_v\rangle = \langle A\rangle_v = \langle\varphi|U^{-1}(v)AU(v)|\varphi\rangle \tag{7.56}$$

Nous nous attendons d'après (7.54) (pour $t = 0$) à ce que les valeurs moyennes des opérateurs position X, impulsion P et vitesse $\dot{X}$ se transforment selon

$$\langle X\rangle \quad\to\quad \langle X\rangle_v = \langle\varphi|U^{-1}(v)XU(v)|\varphi\rangle = \langle X\rangle \tag{7.57}$$

$$\langle \dot{X}\rangle \quad\to\quad \langle \dot{X}\rangle_v = \langle\varphi|U^{-1}(v)\dot{X}U(v)|\varphi\rangle = \langle \dot{X}\rangle + v \tag{7.58}$$

$$\langle P\rangle \quad\to\quad \langle P\rangle_v = \langle\varphi|U^{-1}(v)PU(v)|\varphi\rangle = \langle P\rangle + mv \tag{7.59}$$

L'hypothèse forte[16], même si elle semble naturelle, est en fait (7.58), car $\dot{X}$ est *défini* comme $(\mathrm{i}/\hbar)[H, X]$ (*cf.* (7.18)), et (7.58) conduit à restreindre les hamiltoniens possibles. Comme (7.59) est valable quel que soit $|\varphi\rangle$, on déduit

$$\exp\Big(\mathrm{i}\,\frac{vG}{\hbar}\Big)P\exp\Big(-\mathrm{i}\,\frac{vG}{\hbar}\Big) = P + mvI \tag{7.60}$$

et en faisant tendre v vers zéro

$$[G, P] = -\mathrm{i}\hbar m I$$

16. Voir Brown et Holland [1999] pour une évaluation critique de cette hypothèse.

Il est donc possible de choisir $G = -mX$. D'après le théorème de von Neumann, tout autre choix serait unitairement équivalent.

Considérons maintenant l'opérateur $\dot{X}$ décrivant la vitesse, qui, suivant (7.18) pour $A = X$, est *défini* par

$$\dot{X} = \frac{\mathrm{i}}{\hbar}\,[H, X] \tag{7.61}$$

D'après (7.58),

$$\exp\Big(\mathrm{i}\,\frac{vG}{\hbar}\Big)\dot{X}\exp\Big(-\mathrm{i}\,\frac{vG}{\hbar}\Big) = \dot{X} + vI \tag{7.62}$$

et en retranchant (7.60) (divisé par m) de (7.62)

$$\exp\Big(\mathrm{i}\,\frac{vG}{\hbar}\Big)\Big[\dot{X} - \frac{1}{m}\,P\Big]\exp\Big(-\mathrm{i}\,\frac{vG}{\hbar}\Big) = \dot{X} - \frac{1}{m}\,P \tag{7.63}$$

ce qui implique que l'opérateur $[\dot{X} - P/m]$ commute avec G, et donc avec X. Toujours d'après le théorème de von Neumann, $[\dot{X} - P/m]$ doit être une fonction de X

$$\dot{X} - P/m = \frac{1}{m}\,f(X) \tag{7.64}$$

Dans le cas à une dimension, et en général seulement dans ce cas, on peut éliminer la fonction f par une transformation unitaire. En effet, soit $F(x)$ une primitive de $f(x)$, $F'(x) = f(x)$; considérons la transformation unitaire, qui est une transformation de jauge locale (*cf.* le § 10.3.1)

$$S = \exp\Big(\frac{\mathrm{i}}{\hbar}\,F(X)\Big) \tag{7.65}$$

Dans la transformation unitaire $X' = S^{-1}XS$, X reste évidemment inchangé : $X' = X$. Calculons P'. En utilisant (7.41),

$$[P, S] = -\mathrm{i}\hbar\frac{\partial S}{\partial X} = (-\mathrm{i}\hbar)\Big(\frac{\mathrm{i}}{\hbar}\Big)\,f(X)S = f(X)S$$

d'où l'on déduit

$$S^{-1}PS - P = S^{-1}(PS - SP) = S^{-1}[P, S] = S^{-1}f(X)S = f(X)$$

d'où $P' = S^{-1}PS = P + f(X)$ et d'après (7.64)

$$\dot{X} = \frac{1}{m}\,P'$$

On peut donc toujours choisir comme opérateur impulsion $P = m\dot{X}$: ce choix est unitairement équivalent à tous les autres. Nous allons utiliser ces résultats pour déterminer la forme la plus générale du hamiltonien compatible avec les

lois de transformation galiléennes. Définissons l'opérateur K, qui sera bien sûr la version quantique de l'énergie cinétique, par

$$K = \frac{1}{2}\, m\dot{X}^2 = \frac{P^2}{2m} \tag{7.66}$$

et calculons son commutateur avec X

$$[K, X] = \frac{1}{2m}\,[P^2, X] = -\frac{\mathrm{i}\hbar}{2m}\,\frac{\partial P^2}{\partial P} = -\mathrm{i}\hbar\frac{P}{m} \tag{7.67}$$

En effet, (7.41) implique, en échangeant les rôles de P et de X

$$[X, f(P)] = \mathrm{i}\hbar\,\frac{\partial f(P)}{\partial P}$$

Mais

$$\frac{1}{m}\,P = \dot{X} = \frac{\mathrm{i}}{\hbar}\,[H, X]$$

et en retranchant cette équation de (7.67), on obtient

$$[H - K, X] = 0$$

L'opérateur $(H - K)$ est une fonction de X uniquement, que nous désignerons par $V(X)$, ce qui donne la forme la plus générale du hamiltonien compatible avec l'invariance galiléenne

$$H = K + V(X) = \frac{P^2}{2m} + V(X) \tag{7.68}$$

On justifie ainsi ce que l'on aurait obtenu par le principe de correspondance à partir de l'analogue classique de l'énergie, somme de l'énergie cinétique et de l'énergie potentielle

$$E = \frac{p^2}{2m} + V(x)$$

L'invariance galiléenne est assurée par le fait que le hamiltonien garde la même *forme* après transformation. Si le hamiltonien initial est une fonction de X et P, le hamiltonien transformé est *la même* fonction de $X_v = X$ et $P_v = P + mv$.

- État initial :

$$H = \frac{P^2}{2m} + V(X)$$

- État transformé :

$$H_v = \frac{P_v^2}{2m} + V(X_v) = H + Pv + \frac{1}{2}\, mv^2 + V(X) \tag{7.69}$$

7.4.2 Hamiltonien en dimension $d = 3$

En répétant l'argument de la sous-section précédente dans le cas de trois dimensions d'espace, on arrive sans difficulté à la généralisation de (7.64)

$$\frac{\mathrm{d}\vec{R}}{\mathrm{d}t} = \frac{1}{m}\,\vec{P} - \frac{1}{m}\,\vec{f}(\vec{R}) \tag{7.70}$$

mais on ne peut pas, en général, éliminer $f(\vec{R})$. En effet, il faudrait trouver une transformation unitaire

$$S = \exp\left(\frac{\mathrm{i}}{\hbar}\,F(\vec{R})\right)$$

telle que

$$\vec{f}(\vec{R}) = \vec{\nabla}F(\vec{R})$$

ce qui n'est possible que si $\vec{\nabla} \times \vec{f} = 0$[17]. L'équation (7.70) implique la relation de commutation

$$[\dot{X}_i, X_j] = -\frac{\mathrm{i}\hbar}{m}\,\delta_{ij} \tag{7.71}$$

L'énergie cinétique K est définie par

$$K = \frac{1}{2}\,m\Big(\frac{\mathrm{d}\vec{R}}{\mathrm{d}t}\Big)^2 = \frac{1}{2m}\,(\vec{P} - \vec{f}(\vec{R}))^2 \tag{7.72}$$

On calcule aisément le commutateur de K et de X_i. En effet,

$$[K, X_i] = \frac{1}{2}\,m\sum_j [\dot{X}_j^2, X_i] = \frac{1}{2}\,m\sum_j (\dot{X}_j[\dot{X}_j, X_i] + [\dot{X}_j, X_i]\dot{X}_j) = -\mathrm{i}\hbar\dot{X}_i$$

En comparant les commutateurs

$$[K, X_i] = -\mathrm{i}\hbar\,\dot{X}_i \quad \text{et} \quad [H, X_i] = -\mathrm{i}\hbar\,\dot{X}_i$$

on déduit

$$[H - K, X_i] = 0$$

$(H - K)$ est une fonction de $\vec{R}$ uniquement : $H = K + V(\vec{R})$. Le hamiltonien le plus général compatible avec l'invariance galiléenne est donc de la forme

$$H = \frac{1}{2m}\left(\vec{P} - \vec{f}(\vec{R})\right)^2 + V(\vec{R}) \tag{7.73}$$

Il est important de souligner la différence entre $\vec{P}/m$ et $\mathrm{d}\vec{R}/\mathrm{d}t$: c'est cette dernière quantité qui donne l'énergie cinétique K

$$K = \frac{1}{2m}\left(\frac{\mathrm{d}\vec{R}}{\mathrm{d}t}\right)^2 \neq \frac{\vec{P}^2}{2m}$$

17. Cette condition est nécessaire mais non suffisante dans un domaine qui n'est pas simplement connexe.

On peut maintenant faire le lien avec la physique classique. En mécanique classique, le hamiltonien d'une particule de charge q dans un champ magnétique $\vec{B}(\vec{r},t) = \vec{\nabla} \times \vec{A}(\vec{r},t)$ et dans un champ électrique $\vec{E}(\vec{r},t) = -\partial_t \vec{A}(\vec{r},t) - \vec{\nabla}\overline{V}(\vec{r},t)$ (qui peuvent dépendre du temps) est[18]

$$H_{\rm cl} = \frac{1}{2m}\left(\vec{p} - q\vec{A}(\vec{r},t)\right)^2 + q\overline{V}(\vec{r},t) \tag{7.74}$$

On obtient donc (7.73) en utilisant le principe de correspondance et en identifiant $q\vec{A} = \vec{f}$ et $q\overline{V} = V$. La signification de ce hamiltonien sera examinée de façon plus approfondie au § 10.3.1, quand nous discuterons l'invariance de jauge locale ; la transformation (7.65) et sa généralisation à trois dimensions sont des tranformations de jauge locales. Si $f(\vec{R})$ peut être éliminé par une telle transformation, cela implique que $\vec{B} = 0$. Cependant, il ne faudrait pas en conclure que $\vec{f}$ et $\overline{V}$ doivent nécessairement être identifiés à des potentiels vecteur et scalaire, car $\vec{f}$ et$\overline{V}$ sont des fonctions arbitraires, qui n'ont aucune raison d'obéir aux équations de Maxwell, et la particule n'est pas obligatoirement chargée. Tout ce que nous avons montré est que le hamiltonien classique (7.74) peut être quantifié de façon compatible avec l'invariance de Galilée.

Pour conclure ce chapitre, montrons comment on peut trouver le facteur gyromagnétique pour une particule de spin 1/2 galiléenne (Lévy-Leblond [1967]). Dans l'exercice 8.6.14, on construit explicitement la fonction d'onde transformée dans une transformation de Galilée pour une particule de spin zéro. Le hamiltonien suivant d'une particule libre de spin 1/2

$$H = \frac{1}{2m}(\vec{\sigma}\cdot\vec{P})^2 = \frac{1}{2m}\vec{P}^{\,2} \tag{7.75}$$

agissant dans l'espace de Hilbert produit tensoriel $L^{(2)}(\mathbb{R}^3) \otimes \mathcal{H}_2$, où $\mathcal{H}_2$ est l'espace des états de spin, possède manifestement les mêmes propriétés de covariance par transformation de Galilée que celui d'une particule de spin zéro. Le couplage à un champ magnétique $\vec{P} \to \vec{P} - q\vec{A}$ donne, en utilisant (3.50),

$$H = \frac{1}{2m}\left[\vec{\sigma}\cdot(\vec{P} - q\vec{A})\right]^2 = \frac{1}{2m}(\vec{P} - q\vec{A})^2 - \frac{q\hbar}{2m}(\vec{\sigma}\cdot\vec{B}) \tag{7.76}$$

avec $\vec{B} = \vec{\nabla}\times\vec{A}$; on notera que $\vec{P}$ et $\vec{A}$ ne commutent pas. Le deuxième terme de (7.76) correspond au couplage d'un moment magnétique $\vec{\mu}$ au champ $\vec{B}$, avec

$$\vec{\mu} = \frac{q\hbar}{2m}\vec{\sigma} = \frac{q}{m}\vec{S} = \overline{\gamma}\frac{q}{2m}\vec{S} \tag{7.77}$$

où $\vec{S} = \hbar\vec{\sigma}/2$ est le spin et $\overline{\gamma}$ un facteur égal à 2. Les corrections de théorie quantique des champs invariante de Lorentz, et non de Galilée, font que ce

18. *cf.* Jackson[2001], chapitre 12 et exercice 7.5.8.

facteur est en pratique différent de 2, mais il est très voisin de 2 pour un électron.

Résumons le résultat essentiel obtenu dans ce chapitre. En supposant que les valeurs moyennes des propriétés physiques se transforment dans une symétrie comme les quantités classiques associées, nous avons pu déduire les relations de commutation canoniques et la forme du hamiltonien. Nous n'avons jamais fait appel au principe de correspondance, mais vérifié la compatibilité de ce principe avec nos résultats.

7.5 Exercices

7.5.1 Rotations

1. Soit $\mathcal{R}_{\hat{n}}(\theta)$ la matrice 3×3 représentant une rotation d'angle θ autour de $\hat{n}$. Montrer que $\mathrm{Tr}\,\mathcal{R}_{\hat{n}}(\theta) = 1 + 2\cos\theta$. Suggestion : utiliser (7.29).

2. Écrire explicitement la matrice $\mathcal{R}_{\hat{n}}(\theta)$ à partir de (7.23) en fonction des composantes de $\hat{n}$

$$\hat{n} = (\alpha, \beta, \gamma) \qquad \alpha^2 + \beta^2 + \gamma^2 = 1$$

3. Vérifier explicitement la relation de commutation $[T_x, T_y] = \mathrm{i}T_z$ en utilisant les formes matricielles (7.26).

4. Vérifier que

$$(\vec{T} \cdot \hat{n})^3 = \vec{T} \cdot \hat{n}$$

et en déduire

$$\mathrm{e}^{-\mathrm{i}\theta(\vec{T}\cdot\hat{n})} = I - \mathrm{i}\sin\theta(\vec{T} \cdot \hat{n}) - (1 - \cos\theta)(\vec{T} \cdot \hat{n})^2$$

Comparer avec (7.23).

7.5.2 Rotations et $SU(2)$

Le groupe $SU(2)$ est le groupe des matrices 2×2 unitaires et de déterminant un.

1. Montrer que si $U \in SU(2)$, alors U est de la forme

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \qquad |a|^2 + |b|^2 = 1$$

2. Montrer qu'au voisinage de l'identité, on peut écrire

$$U = I - \mathrm{i}\tau \quad \text{avec} \quad \tau = \tau^\dagger$$

et que τ s'exprime en fonction des matrices de Pauli

$$\tau = \frac{1}{2}\sum_{i=1}^{3} \theta_i \sigma_i \qquad \theta_i \to 0$$

3. On pose $\theta = (\sum_i \theta_i^2)^{1/2}$ et $\theta_i = \theta\,\hat{n}_i$, où $\hat{n}$ est un vecteur unitaire. Supposant maintenant les θ_i finis, on définit $U_{\hat{n}}(\theta)$ par

$$U_{\hat{n}}(\theta) = \lim_{N\to\infty}\left[U_{\hat{n}}\left(\frac{\theta}{N}\right)\right]^N$$

Montrer que

$$U_{\hat{n}}(\theta) = \mathrm{e}^{-\mathrm{i}\theta\vec{\sigma}\cdot\hat{n}/2}$$

Inversement, toute matrice de $SU(2)$ est de cette forme (exercice 3.3.6).

4. Soit $\vec{V}$ un vecteur de $\mathbb{R}^3$ et $\mathcal{V}$ la matrice hermitienne de trace nulle

$$\mathcal{V} = \begin{pmatrix} V_z & V_x - \mathrm{i}V_y \\ V_x + \mathrm{i}V_y & -V_z \end{pmatrix} = \vec{\sigma}\cdot\vec{V}$$

Quel est le déterminant de $\mathcal{V}$? Soit $\mathcal{W}$ la matrice

$$\mathcal{W} = U\mathcal{V}U^{-1}$$

Montrer que $\mathcal{W}$ est de la forme $\vec{\sigma}\cdot\vec{W}$ et que $\vec{W}$ se déduit de $\vec{V}$ par une rotation. A-t-on entièrement prouvé cette propriété à ce stade ?

5. On définit $\vec{V}(\theta)$ par

$$\vec{\sigma}\cdot\vec{V}(\theta) = U_{\hat{n}}(\theta)\,[\vec{\sigma}\cdot\vec{V}]\,U_{\hat{n}}^{-1}(\theta) \qquad \vec{V}(\theta = 0) = \vec{V}$$

Montrer que

$$\frac{\mathrm{d}\vec{V}(\theta)}{\mathrm{d}\theta} = \hat{n}\times\vec{V}(\theta)$$

En déduire que $\vec{V}(\theta)$ s'obtient à partir de $\vec{V}$ par une rotation d'angle θ autour de $\hat{n}$. Ce résultat établit une correspondance entre les matrices $\mathcal{R}_{\hat{n}}(\theta)$ de $SO(3)$ et les matrices $U_{\hat{n}}(\theta)$ de $SU(2)$. Cette correspondance est-elle biunivoque ?

7.5.3 Relations de commutation entre l'impulsion et le moment angulaire

Cet exercice donne une autre démonstration des relations de commutation (7.34) entre l'impulsion et le moment angulaire, si l'on choisit l'opérateur vectoriel $\vec{V} = \vec{P}$. Soit $\mathcal{T}_y(a)$, une translation de a parallèle à Oy

$$\mathcal{T}_y(a)\vec{r} = \vec{r} + a\hat{y}$$

Si $\mathcal{R}_x(\theta)$ est une rotation d'angle θ autour de Ox, montrer que

$$\mathcal{R}_x(\theta)\,\mathcal{T}_y(a)\,\mathcal{R}_x(-\theta)$$

est une translation le long d'un axe que l'on déterminera. En déduire la relation de commutation

$$[J_x, P_y] = \mathrm{i}\hbar P_z$$

7.5.4 Algèbre de Lie d'un groupe continu

On considère un groupe $\mathcal{G}$ dont les élements g sont paramétrés par N coordonnées $\theta_a,\ a = 1,\ \ldots,\ N$, $g(\theta_a = 0)$ étant l'élément neutre du groupe. Les variables θ_a sont notées collectivement θ : $\theta = \{\theta_a\}$. La loi de composition est donnée par une fonction f indéfiniment différentiable

$$g(\overline{\theta})g(\theta) = g(f(\overline{\theta}, \theta))$$

À nouveau, f est une notation collective pour l'ensemble de N fonctions f : $f(\overline{\theta}, \theta) = \{f_a(\overline{\theta}_b, \theta_c)\}$. Soit un ensemble de matrices unitaires $U(\theta_a)$ dont la loi de multiplication est

$$U(\overline{\theta})U(\theta) = U(f(\overline{\theta}, \theta))$$

Les matrice $U(\theta)$ forment donc une représentation du groupe $\mathcal{G}$: voir (7.12).

1. Montrer que $f_a(\overline{\theta}, \theta = 0) = \overline{\theta}_a$ et que $f_a(\overline{\theta} = 0, \theta) = \theta_a$. Montrer que, pour $\overline{\theta}, \theta \to 0$, $f_a(\overline{\theta}, \theta)$ est de la forme

$$f_a(\overline{\theta}, \theta) = \theta_a + \overline{\theta}_a + f_{abc}\overline{\theta}_b\theta_c + O(\theta^3, \theta^2\,\overline{\theta}, \theta\,\overline{\theta}^2, \overline{\theta}^3)$$

où nous avons utilisé la convention de sommation sur les indices répétés

$$f_{abc}\overline{\theta}_b\theta_c = \sum_{b,c} f_{abc}\overline{\theta}_b\theta_c$$

2. Dans le voisinage de $U(\theta) = I$, on écrit un développement de $U(\theta)$ pour $\theta \to 0$

$$U(\theta) = I - \mathrm{i}\theta_a T_a - \frac{1}{2}\,\theta_b\theta_c T_{bc} + O(\theta)^3$$

Effectuer le produit $U(\overline{\theta})U(\theta)$ à l'ordre $(\overline{\theta}^2, \theta^2)$ et montrer que l'égalité

$$U(\overline{\theta})U(\theta) = U(f(\overline{\theta}, \theta))$$

pour les termes en $\overline{\theta}_a\theta_b$ implique que

$$T_{bc} = T_c T_b - \mathrm{i}f_{abc}\,T_a$$

Utilisant la symétrie de T_{bc}, en déduire

$$[T_b, T_c] = \mathrm{i}C_{abc}\,T_a$$

avec $C_{abc} = -C_{acb}$. Exprimer C_{abc} en fonction de f_{abc}. Les relations de commutation précédentes constituent l'algèbre de Lie du groupe défini par la loi de composition $f(\overline{\theta}, \theta)$.

7.5.5 Règle de somme de Thomas-Reiche-Kuhn

Soit une particule de masse m dans un potentiel $V(\vec{r})$. Le hamiltonien est

$$H = \frac{\vec{P}^2}{2m} + V(\vec{R})$$

Soit $|\varphi_n\rangle$ un ensemble complet de vecteurs propres de H

$$H|\varphi_n\rangle = E_n|\varphi_n\rangle \qquad \sum_n |\varphi_n\rangle\langle\varphi_n| = I$$

et $|\varphi_0\rangle$ un état lié, donc normalisable, d'énergie E_0. On pose

$$\langle\varphi_n|X|\varphi_0\rangle = X_{n0}$$

1. Démontrer la relation de commutation

$$[[H, X], X] = -\frac{\hbar^2}{m}$$

2. En déduire

$$\sum_n \frac{2m|X_{n0}|^2}{\hbar^2}(E_n - E_0) = 1$$

7.5.6 Centre de masse et masse réduite

Soit deux particules de masses m_1 et m_2 se déplaçant sur une droite. On note X_1 et X_2 leurs opérateurs position, P_1 et P_2 leurs opérateurs impulsion. Les opérateurs impulsion et position de deux particules différentes commutent. On définit les opérateurs X et P

$$X = \frac{m_1 X_1 + m_2 X_2}{m_1 + m_2} \qquad P = P_1 + P_2$$

et $\tilde{X}$ et $\tilde{P}$

$$\tilde{X} = X_1 - X_2 \qquad \tilde{P} = \frac{m_2 P_1 - m_1 P_2}{m_1 + m_2}$$

1. Calculer les commutateurs $[X, P]$ et $[\tilde{X}, \tilde{P}]$ et montrer que

$$[X, \tilde{P}] = [\tilde{X}, P] = 0$$

2. Écrire le hamiltonien

$$H = \frac{P_1^2}{2m_1} + \frac{P_2^2}{2m_2} + V(X_1 - X_2)$$

en fonction des opérateurs $X, P, \tilde{X}, \tilde{P}$. En conclure que, comme en mécanique classique, on peut séparer le mouvement du centre de masse et de la particule relative de masse réduite $\mu = m_1 m_2/(m_1 + m_2)$. Généraliser à trois dimensions.

3. L'exemple suivant d'état intriqué a été utilisé dans l'article d'Einstein *et al.* [1935]. La fonction d'onde de deux particules est écrite

$$\psi(x_1, x_2; p_1, p_2) = \delta(x_1 - x_2 - L)\,\delta(p_1 + p_2)$$

où L est une longueur constante. Pourquoi est-il possible d'écrire une telle fonction d'onde ? Quelle est son interprétation physique ? La mesure de x_1 détermine x_2, celle de p_1 détermine p_2. Développer l'analogie avec l'exemple du 11.2.1.

7.5.7 Transformation de Galilée

1. Soit $W(a,v)$ le produit d'une transformation de Galilée et d'une translation à une dimension

$$W(a,v) = \exp\Big(-\mathrm{i}\frac{Pa}{\hbar}\Big)\,\exp\Big(\mathrm{i}\frac{mvX}{\hbar}\Big)$$

Montrer que

$$W(a_1,v_1)W(a_2,v_2) = \exp\Big(-\mathrm{i}\frac{mv_1a_2}{\hbar}\Big)W(a_1+a_2, v_1+v_2)$$

2. Calculer

$$W(a,v)W(-a,-v)$$

et en déduire que l'on doit utiliser des représentations projectives pour le groupe de Galilée.

7.5.8 Hamiltonien dans un champ magnétique

1. On se donne un champ électromagnétique $(\vec{E}, \vec{B})$ dérivant du potentiel $(\overline{V}, \vec{A})$

$$E_i = -\partial_i\overline{V} - \partial_t A_i \qquad B_i = \sum_{jk}\varepsilon_{ijk}\partial_j B_k = (\vec{\nabla}\times\vec{A})_i$$

On se propose de montrer que les équations du mouvement provenant du lagrangien suivant d'une particule chargée de charge q

$$\mathcal{L} = \frac{1}{2}m\left(\frac{\mathrm{d}\vec{r}}{\mathrm{d}t}\right)^2 - q\overline{V}(\vec{r},t) + q\vec{A}(\vec{r},t)\cdot\frac{\mathrm{d}\vec{r}}{\mathrm{d}t}$$

sont identiques à celles données par la loi de Lorentz. Calculer $\partial\mathcal{L}/\partial x_i$, $\partial\mathcal{L}/\partial\dot{x}_i$ et montrer que

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_i} = m\ddot{x}_i + q\sum_j(\partial_j A_i)\dot{x}_j$$

Montrer que les équations d'Euler-Lagrange

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_i} - \frac{\partial\mathcal{L}}{\partial x_i} = 0$$

conduisent à la loi de Lorentz non relativiste ($\vec{v} = \mathrm{d}\vec{r}/\mathrm{d}t$)

$$m\ddot{x}_i = q[E_i + (\vec{v} \times \vec{B})_i]$$

2. Partant du moment conjugué p_k de x_k

$$p_k = \frac{\partial \mathcal{L}}{\partial \dot{x}_k} = m\dot{x}_k + qA_k$$

montrer que le hamiltonien classique $H_{\mathrm{cl}} = \sum_k p_k \dot{x}_k - L$ s'écrit

$$H_{\mathrm{cl}} = \frac{1}{2m}\,(\vec{p} - q\vec{A})^2 + q\overline{V}$$

3. Retrouver la loi de Lorentz à partir des équations de Hamilton

$$\dot{x}_k = \frac{\partial H_{\mathrm{cl}}}{\partial p_k} \qquad \dot{p}_k = -\frac{\partial H_{\mathrm{cl}}}{\partial x_k}$$

7.6 Bibliographie

On trouvera des compléments utiles sur les symétries en physique quantique dans Jauch [1968], chapitres 9 et 10 et dans Merzbacher [1970], chapitre 16. Le chapitre 2 de Weinberg [1995] contient également un excellent résumé de toutes les notions de base. Les relations de commutation canoniques et l'invariance galiléenne sont exposées dans Jauch [1968], chapitres 12 et 13. Il existe de nombreux livres consacrés à l'utilisation de la théorie des groupes en mécanique quantique, parmi lesquels on peut citer Tinkham [1964].

Chapitre 8

Mécanique ondulatoire

Dans ce chapitre, nous allons étudier une réalisation particulière de la mécanique quantique d'une grande importance pratique : il s'agit de la *mécanique ondulatoire*, qui est adaptée à la description du mouvement d'une particule quantique[1] dans l'espace à trois dimensions $\mathbb{R}^3$. C'est cette réalisation qui sert d'introduction aux fondements de la mécanique quantique dans la plupart des manuels. Elle consiste à prendre comme base de $\mathcal{H}$ les "vecteurs propres[2]" $|\vec{r}\,\rangle$ de l'opérateur position $\vec{R}$: en d'autres termes, on choisit une base où l'opérateur position est diagonal. En mécanique ondulatoire, un vecteur d'état peut être identifié à un élément $\varphi(\vec{r}\,)$ de l'espace de Hilbert $L^2_{\vec{r}}(\mathbb{R}^3)$ des fonctions de carré sommable sur l'espace à trois dimensions $\mathbb{R}^3$. Ce vecteur d'état est appelé *fonction d'onde*, et nous verrons que cette fonction d'onde s'identifie à l'amplitude de probabilité $\langle\vec{r}\,|\varphi\rangle$ de trouver la particule dans l'état $|\varphi\rangle$ localisée à la position $\vec{r}$. Cette fonction d'onde est normalisée par la condition de sommabilité (6.10)

$$\int_{-\infty}^{\infty} \mathrm{d}^3r\, |\varphi(\vec{r})|^2 = 1 \tag{8.1}$$

Étant donné le rôle symétrique joué par les opérateurs position et impulsion, on pourrait aussi bien utiliser les vecteurs propres de $\vec{P}$ et les "fonctions d'onde de l'espace des impulsions" $\tilde{\varphi}(\vec{p}) = \langle\vec{p}\,|\varphi\rangle$, dont nous verrons que ce sont les transformées de Fourier des $\varphi(\vec{r})$. Après avoir examiné les principales propriétés des fonctions d'onde, nous étudierons plusieurs applications : états liés, diffusion, potentiel périodique. Ces applications seront d'abord traitées dans le cas plus simple à une dimension. La généralisation à trois dimensions nous permettra de discuter la notion importante de densité d'états et son utilisation dans la "règle d'or de Fermi".

1. Ou de plusieurs particules : voir la généralisation au chapitre 14.

2. Ainsi que nous l'avons vu au § 6.3.1, ces objets ne sont pas des vecteurs de l'espace de Hilbert, ce que nous avons souligné par des guillemets. Cependant, comme nous allons faire par la suite un usage intensif de ces "vecteurs", nous supprimerons ces guillemets afin d'alléger l'écriture.

8.1 Diagonalisation de X et de P ; fonctions d'onde

8.1.1 Diagonalisation de X

Nous nous proposons d'étudier le mouvement d'une particule quantique en nous restreignant pour le moment au mouvement sur la droite réelle $\mathbb{R}$: la particule se déplace sur cette droite entre $-\infty$ et $+\infty$. Les grandeurs physiques pertinentes sont *a priori* la position et l'impulsion de cette particule, représentées mathématiquement par des opérateurs X et P dont nous avons établi les propriétés dans la section 7.3. Nous allons d'abord examiner les vecteurs propres de X en partant de la relation de commutation canonique entre X et P sous la forme (7.40)

$$\exp\Big(\mathrm{i}\frac{Pa}{\hbar}\Big)\, X \,\exp\Big(-\mathrm{i}\frac{Pa}{\hbar}\Big) = X + aI \tag{8.2}$$

Montrons d'abord que le spectre de X est continu : soit $|x\rangle$ un vecteur propre de X

$$X|x\rangle = x|x\rangle \tag{8.3}$$

et examinons l'action de X sur le vecteur $\exp(-\mathrm{i}Pa/\hbar)|x\rangle$

$$\begin{aligned} X\Big[\exp\Big(-\mathrm{i}\frac{Pa}{\hbar}\Big)|x\rangle\Big] &= \exp\Big(-\mathrm{i}\frac{Pa}{\hbar}\Big)(X+aI)|x\rangle \\ &= (x+a)\Big[\exp\Big(-\mathrm{i}\frac{Pa}{\hbar}\Big)|x\rangle\Big] \end{aligned} \tag{8.4}$$

Nous avons utilisé la relation de commutation (8.2) et la définition (8.3) du vecteur propre $|x\rangle$. Le vecteur $\exp(-\mathrm{i}Pa/\hbar)|x\rangle$ est vecteur propre de X avec la valeur propre $(x+a)$, et comme a est arbitraire, cela montre que toutes les valeurs réelles de x entre $-\infty$ et $+\infty$ sont valeurs propres de X. On prouve ainsi que le spectre de x est continu, et que, par conséquent, la normalisation doit s'écrire suivant (6.34) avec des fonctions delta de Dirac

$$\langle x'|x\rangle = \delta(x-x') \tag{8.5}$$

Compte tenu de la discussion du § 7.3.1, le résultat (8.4) qui peut s'écrire

$$\exp\Big(-\mathrm{i}\frac{Pa}{\hbar}\Big)|x\rangle = |x+a\rangle$$

ne doit pas surprendre, car $\exp(-\mathrm{i}Pa/\hbar)$ est l'opérateur de translation de a qui transforme l'état localisé en x, $|x\rangle$, en l'état $|x+a\rangle$ localisé en $(x+a)$. Le vecteur $|x+a\rangle$ vérifie une condition de normalisation analogue à (8.5), étant donné que l'opérateur $\exp(-\mathrm{i}Pa/\hbar)$ est unitaire. On peut, si on le souhaite, fixer la phase des vecteurs de base $|x\rangle$ par la condition

$$|x\rangle = \exp\Big(-\mathrm{i}\frac{Px}{\hbar}\Big)|x=0\rangle \tag{8.6}$$

Revenons sur l'interprétation physique : que représente exactement le vecteur $|x\rangle$? D'après les postulats du chapitre 4, $|x\rangle$ représente un état où la position de la particule est connue avec une précision absolue : la particule est localisée *exactement* au point x sur la droite réelle. Mais, en mécanique quantique un tel état est impossible à réaliser physiquement : comme nous allons bientôt le voir, un tel état a toutes les impulsions possibles avec la même probabilité entre $p = -\infty$ et $p = +\infty$. À la propriété mathématique : "$|x\rangle$ n'est pas un vecteur de l'espace de Hilbert" correspond la propriété physique : "$|x\rangle$ n'est pas un état physiquement réalisable". Les états physiquement réalisables sont toujours représentés par des "vrais" vecteurs de $\mathcal{H}$, c'est-à-dire des vecteurs normalisables.

Nous avons admis implicitement que les valeurs propres x de X étaient non dégénérées. Bien sûr, ce n'est pas nécessairement le cas : par exemple, la particule pourrait avoir un spin 1/2, auquel cas il faudrait préciser si la particule se trouve dans un état de spin up $|+\rangle$ ou de spin down $|-\rangle$: chaque valeur propre de X serait alors doublement dégénérée. Dans ces conditions, l'espace de Hilbert des états serait le produit tensoriel $L_x^{(2)}(\mathbb{R}) \otimes \mathcal{H}_2$ de l'espace des états de position $L_x^{(2)}(\mathbb{R})$ et de l'espace à deux dimensions des états de spin $\mathcal{H}_2$: une base de cet espace serait par exemple formée des états $|x \otimes +\rangle$ et $|x \otimes -\rangle$ avec

$$(X \otimes \sigma_z)|x \otimes \pm\rangle = \pm x|x \otimes \pm\rangle$$

Bien que l'utilisation de vecteurs propres qui ne soient pas des véritables éléments de $\mathcal{H}$ soit mathématiquement contestable, elle est extrêmement commode et nous en servirons par la suite sans précautions particulières. Nous généraliserons aussi la notion d'élément de matrice : comme l'opérateur X est diagonal dans la base $|x\rangle$, nous pourrons écrire les "éléments de matrice" de X

$$\langle x'|X|x\rangle = x\langle x'|x\rangle = x\,\delta(x - x') \tag{8.7}$$

et plus généralement ceux d'une fonction $F(X)$

$$\langle x'|F(X)|x\rangle = F(x)\langle x'|x\rangle = F(x)\,\delta(x - x') \tag{8.8}$$

La relation de fermeture (6.36) s'écrit

$$\int_{-\infty}^{\infty} |x\rangle\,\mathrm{d}x\,\langle x| = I \tag{8.9}$$

Le projecteur $\mathcal{P}[a, b]$ sur le sous-espace des valeurs propres de X dans l'intervalle $[a, b]$ s'obtient en restreignant l'intégration sur x à cet intervalle

$$\mathcal{P}[a, b] = \int_a^b |x\rangle\,\mathrm{d}x\,\langle x| \tag{8.10}$$

Cette expression généralise celle que l'on écrirait dans un espace de dimension finie : si Δ est le sous-espace d'un ensemble de valeurs propres d'un opérateur

hermitien A, le projecteur $\mathcal{P}(\Delta)$ sur ce sous-espace est

$$\mathcal{P}(\Delta) = \sum_{n\in\Delta} |n\rangle\langle n|$$

8.1.2 Réalisation dans $L^{(2)}_x(\mathbb{R})$

Nous allons maintenant faire le lien entre le formalisme de Dirac que nous venons d'expliciter dans la base où X est diagonal et la réalisation donnée au § 7.3.2 des opérateurs X et P comme opérateurs agissant dans l'espace $L^{(2)}(\mathbb{R})$ des fonctions de carré sommable sur $\mathbb{R}$. Soit $|\varphi\rangle$ un vecteur unitaire de $\mathcal{H}$ représentant un état physique. En utilisant la relation de fermeture (8.9), nous pouvons décomposer $|\varphi\rangle$ dans la base $|x\rangle$

$$|\varphi\rangle = \int_{-\infty}^{\infty} |x\rangle \,\mathrm{d}x\, \langle x|\varphi\rangle \tag{8.11}$$

où $\langle x|\varphi\rangle$ est donc une composante de $|\varphi\rangle$ dans la base $|x\rangle$, ou en termes physiques, l'amplitude de probabilité de trouver la particule localisée au point x. Examinons les éléments de matrice des opérateurs X et $\exp(-\mathrm{i}Pa/\hbar)$

$$\langle x|\big[X|\varphi\rangle\big] = \langle Xx|\varphi\rangle = x\,\langle x|\varphi\rangle = x\,\varphi(x) \tag{8.12}$$

$$\langle x\Big|\Big[\exp\Big(-\mathrm{i}\frac{Pa}{\hbar}\Big)\Big|\varphi\rangle\Big] = \langle x-a|\varphi\rangle = \varphi(x-a) \tag{8.13}$$

Ces équations montrent que $\langle x|\varphi\rangle$ peut être identifié à une fonction $\varphi(x)$ de $L^{(2)}_x(\mathbb{R})$ telle que l'action des opérateurs X et P soit donnée par (7.44). En effet, l'équation (8.12) est

$$\boxed{\big[X\varphi\big](x) = x\varphi(x)} \tag{8.14}$$

tandis que (8.13) s'écrit

$$\Big[\exp\Big(-\mathrm{i}\frac{Pa}{\hbar}\Big)\varphi\Big](x) = \varphi(x-a) \tag{8.15}$$

et en développant au premier ordre en a

$$\boxed{\big[P\varphi\big](x) = -\mathrm{i}\hbar\,\frac{\partial\varphi}{\partial x}} \tag{8.16}$$

Nous retrouvons l'action des opérateurs X et P telle qu'elle avait été définie au § 7.3.2. Vérifions que le produit scalaire est correctement donné par (8.9) en utilisant la relation de fermeture (8.9)

$$\langle\chi|\varphi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\langle\chi|x\rangle\langle x|\varphi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\chi^*(x)\varphi(x) \tag{8.17}$$

La fonction $\varphi(x-a)$ dans (8.15) est bien la fonction $\varphi(x)$ translatée de $+a$, et non de $-a$! Si par exemple $\varphi(x)$ a un maximum à $x = x_0$, $\varphi(x-a)$ a un maximum à $x - a = x_0$, c'est-à-dire à $x = x_0 + a$ (figure 8.1). Soulignons que le choix $\varphi_a(x) = \varphi(x-a)$ pour la fonction d'onde translatée n'est bien sûr pas unique. La fonction

$$\varphi'_a(x) = \mathrm{e}^{\mathrm{i}\theta(x)}\varphi(x-a)$$

se déduit de $\varphi(x-a)$ par une transformation de jauge locale (7.65). Le choix $\varphi(x-a)$ est lié à celui du générateur infinitésimal des translations, et le changement de phase $\varphi_a(x) \to \varphi'_a(x)$ serait obtenu en utilisant un générateur infinitésimal déduit de (9.16) par la transformation de jauge locale

$$P' = \mathrm{e}^{\,\mathrm{i}\theta(x)}\left(\mathrm{i}\hbar\frac{\partial}{\partial x}\right)\mathrm{e}^{-\mathrm{i}\theta(x)}$$

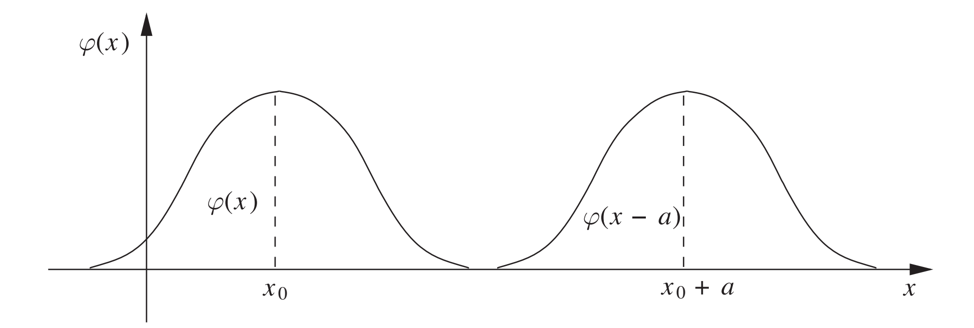


FIG. 8.1 – Translation de a de la fonction d'onde d'une particule localisée en x_0.

En résumé, l'état physique d'une particule se déplaçant sur l'axe des x est décrit par une fonction d'onde normalisée $\varphi(x)$ appartenant à $L^{(2)}_x(\mathbb{R})$

$$\int_{-\infty}^{\infty} \mathrm{d}x\, |\varphi(x)|^2 = 1 \tag{8.18}$$

qui s'interprète physiquement comme l'amplitude de probabilité $\langle x|\varphi\rangle$ de trouver la particule localisée au point x. L'action des opérateurs position X et impulsion P sur $\varphi(x)$ est donnée par (8.14) et (8.16). Le module carré

$$|\varphi(x)|^2 = |\langle x|\varphi\rangle|^2$$

est appelé *probabilité de présence* de la particule au point x : en fait, il s'agit d'une *densité de probabilité*, dans ce cas une probabilité par unité de longueur. La probabilité $\mathsf{p}([a,b])$ de trouver la particule localisée dans l'intervalle $[a,b]$ est d'après (8.10)

$$\mathsf{p}([a,b]) = \langle\varphi|\mathcal{P}[a,b]|\varphi\rangle = \int_a^b \mathrm{d}x\, |\varphi(x)|^2 \tag{8.19}$$

Cette probabilité est normalisée à un par construction, puisque $\langle\varphi|\varphi\rangle = 1$, ce que l'on vérifie également sur (8.18). Si l'on considère un intervalle $[x, x + \mathrm{d}x]$ infinitésimal, $|\varphi(x)|^2\mathrm{d}x$ est la probabilité de trouver la particule dans cet intervalle.

Lorsque la particule possède des degrés de liberté supplémentaires, par exemple un spin 1/2, on pourra décrire son état quantique en utilisant les fonctions d'onde $\varphi_\pm(x)$

$$\varphi_+(x) = \langle x \otimes +|\varphi\rangle \qquad \varphi_-(x) = \langle x \otimes -|\varphi\rangle$$

Nous venons de définir ce que l'on appelle habituellement "la mécanique ondulatoire en représentation x" : nous avons choisi de partir de la base $|x\rangle$ où l'opérateur position est diagonal. Étant donné le rôle symétrique de X et P, nous aurions aussi bien pu partir d'une base où P est diagonal, c'est-à-dire définir "la mécanique ondulatoire en représentation p". La sous-section suivante est consacrée à cette représentation et à son lien avec la représentation x.

8.1.3 Réalisation dans $L_p^{(2)}(\mathbb{R})$

Soit $|p\rangle$ un vecteur propre de P

$$P|p\rangle = p|p\rangle \tag{8.20}$$

Nous allons d'abord déterminer les fonctions d'onde correspondantes $\chi_p(x) = \langle x|p\rangle$. Nous pouvons écrire en représentation x

$$\langle x|\big[P|p\rangle\big] = p\langle x|p\rangle = p\,\chi_p(x) = -\mathrm{i}\hbar\,\frac{\partial}{\partial x}\,\chi_p(x) \tag{8.21}$$

Nous avons utilisé (8.16) pour obtenir la dernière égalité de l'équation précédente. Quel que soit p dans l'intervalle $]-\infty, +\infty[$, l'équation différentielle

$$-\mathrm{i}\hbar\,\frac{\partial}{\partial x}\,\chi_p(x) = p\,\chi_p(x)$$

a pour solution

$$\boxed{\chi_p(x) = \langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\,\mathrm{e}^{\mathrm{i}px/\hbar}} \tag{8.22}$$

ce qui montre que le spectre de P est continu, tout comme celui de x. Le facteur de normalisation $(2\pi\hbar)^{-1/2}$ dans (8.22) a été choisi de telle sorte que $\chi_p(x)$ soit normalisé par un delta de Dirac

$$\int_{-\infty}^{\infty} \mathrm{d}x\,\chi_{p'}^*(x)\chi_p(x) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} \mathrm{d}x\,\exp\left[\mathrm{i}\,\frac{(p-p')x}{\hbar}\right] = \delta(p-p') \tag{8.23}$$

tandis que la relation de fermeture s'écrit

$$\int_{-\infty}^{\infty} \mathrm{d}p\,\chi_p(x)\chi_p^*(x') = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} \mathrm{d}p\,\exp\left[\mathrm{i}\,\frac{p(x-x')}{\hbar}\right] = \delta(x-x') \tag{8.24}$$

On aurait également pu partir de la relation de fermeture sous la forme

$$\int_{-\infty}^{\infty} |p\rangle \, \mathrm{d}p \, \langle p| = I \tag{8.25}$$

pour écrire

$$\int_{-\infty}^{\infty} \langle x'|p\rangle \, \mathrm{d}p \, \langle p|x\rangle = \langle x'|I|x\rangle = \delta(x - x')$$

et démontrer ainsi (8.24).

Si $|\varphi\rangle$ est le vecteur d'état d'une particule, la "fonction d'onde en représentation p" sera $\tilde{\varphi}(p) = \langle p|\varphi\rangle$. Cette fonction d'onde en représentation p n'est autre que la transformée de Fourier de la fonction d'onde $\varphi(x) = \langle x|\varphi\rangle$ de la représentation x. En effet, en utilisant la relation de fermeture (8.9) et la définition (8.22), nous obtenons

$$\boxed{\tilde{\varphi}(p) = \langle p|\varphi\rangle = \int_{-\infty}^{\infty} \langle p|x\rangle \, \mathrm{d}x \, \langle x|\varphi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \mathrm{d}x \, \mathrm{e}^{-\mathrm{i}px/\hbar} \, \varphi(x)} \tag{8.26}$$

et inversement

$$\boxed{\varphi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \mathrm{d}p \, \mathrm{e}^{\mathrm{i}px/\hbar} \, \tilde{\varphi}(p)} \tag{8.27}$$

L'action des opérateurs X et P en représentation p s'obtient sans difficulté

$$\big[X\tilde{\varphi}\big](p) = \mathrm{i}\hbar \frac{\partial}{\partial p} \tilde{\varphi}(p) \tag{8.28}$$

$$\big[P\tilde{\varphi}\big](p) = p \, \tilde{\varphi}(p) \tag{8.29}$$

Une formule analogue à (8.19) est valable dans l'espace des impulsions : la probabilité $\mathsf{p}([k,q])$ pour que la particule ait son impulsion dans l'intervalle $[k,q]$ est

$$\mathsf{p}([k,q]) = \int_{k}^{q} \mathrm{d}p \, |\tilde{\varphi}(p)|^2 \tag{8.30}$$

$|\tilde{\varphi}(p)|^2$ est une densité de probabilité dans l'espace des impulsions.

8.1.4 Inégalités de Heisenberg

Partons de la représentation de Fourier (8.27) de la fonction d'onde $\varphi(x)$ d'un état physique. La transformée de Fourier $\tilde{\varphi}(p)$ vérifie, tout comme $\varphi(x)$, la condition de normalisation

$$\int_{-\infty}^{\infty} \mathrm{d}p \, |\tilde{\varphi}(p)|^2 = 1 \tag{8.31}$$

On appelle souvent un tel état physique un *paquet d'ondes*, car c'est d'après (8.27) une superposition d'ondes planes. La position moyenne $\langle X\rangle$ et l'impulsion moyenne $\langle P\rangle$ se calculent en introduisant deux fois les relations de fermeture (8.9) et (8.25)[3]

$$\langle X\rangle = \langle\varphi|X|\varphi\rangle = \int \mathrm{d}x\,\mathrm{d}x'\langle\varphi|x\rangle\langle x|X|x'\rangle\langle x'|\varphi\rangle = \int_{-\infty}^{\infty}\mathrm{d}x\,x|\varphi(x)|^2 \tag{8.32}$$

$$\langle P\rangle = \langle\varphi|P|\varphi\rangle = \int \mathrm{d}p\,\mathrm{d}p'\langle\varphi|p\rangle\langle p|P|p'\rangle\langle p'|\varphi\rangle = \int_{-\infty}^{\infty}\mathrm{d}p\,p|\tilde{\varphi}(p)|^2 \tag{8.33}$$

Nous avons aussi utilisé (8.7) et une équation analogue dans l'espace des impulsions. Les dispersions ΔX et ΔP sont données par un calcul similaire

$$(\Delta X)^2 = \langle\varphi|(X-\langle X\rangle)^2|\varphi\rangle = \int_{-\infty}^{\infty}\mathrm{d}x\,(x-\langle X\rangle)^2|\varphi(x)|^2 \tag{8.34}$$

$$(\Delta P)^2 = \langle\varphi|(P-\langle P\rangle)^2|\varphi\rangle = \int_{-\infty}^{\infty}\mathrm{d}p\,(p-\langle P\rangle)^2|\tilde{\varphi}(p)|^2 \tag{8.35}$$

D'après la démonstration générale du § 4.1.3, ces dispersions vérifient l'inégalité de Heisenberg

$$\boxed{\Delta x\,\Delta p \geq \frac{1}{2}\,\hbar} \tag{8.36}$$

où nous avons utilisé la notation usuelle $\Delta x\,\Delta p$ au lieu de $\Delta X\,\Delta P$. Une démonstration directe de (8.36) est proposée dans l'exercice 8.7.1.

Pour conclure cette sous-section, montrons l'intérêt de l'inégalité de Heisenberg (8.36) utilisée comme outil heuristique en estimant l'énergie de l'état fondamental de l'atome d'hydrogène (voir le § 1.5.2). Si l'électron décrit une orbite circulaire de rayon r avec une impulsion $p = mv$, son énergie classique est

$$E = \frac{p^2}{2m} - \frac{e^2}{r} \tag{8.37}$$

En physique classique, le rayon de l'orbite de l'électron tend vers zéro ("l'électron tombe sur le noyau"), ce phénomène étant accompagné de l'émission de rayonnement électromagnétique : en effet, en physique classique, l'énergie de l'orbite circulaire $E = -e^2/(2r)$ n'est pas bornée inférieurement et rien ne s'oppose à ce que le rayon de l'orbite devienne arbitrairement petit. La décroissance de l'énergie de l'orbite est compensée par l'émission dans l'espace d'énergie sous forme de rayonnement électromagnétique, ce qui assure la conservation de l'énergie. Mais sur une orbite de rayon r, la dispersion Δx de la position suivant l'axe des x est de l'ordre de r, ce qui fait que la dispersion

3. Les notations explicites seraient $\langle X\rangle_\varphi$ et $\langle P\rangle_\varphi$; nous avons supprimé l'indice φ pour alléger l'écriture.

sur l'impulsion est au moins de $\sim \hbar/\Delta x = \hbar/r$. Nous pouvons en déduire $rp \sim \hbar$ et l'expression de l'énergie (8.37) devient

$$E \sim \frac{\hbar^2}{2mr^2} - \frac{e^2}{r}$$

Cherchons le minimum de E

$$\frac{\mathrm{d}E}{\mathrm{d}r} \sim -\frac{\hbar^2}{mr^3} + \frac{e^2}{r^2} = 0$$

ce qui donne un minimum pour

$$r = a_0 = \frac{\hbar^2}{me^2} \tag{8.38}$$

soit précisément le rayon de Bohr (1.41) de l'atome d'hydrogène ! Naturellement, le fait que l'on obtienne exactement a_0 dans ce calcul d'ordre de grandeur est un hasard heureux, qui nous permet de retrouver l'énergie de l'état fondamental (1.42)

$$E_0 = -\frac{e^2}{2a_0} = -\frac{me^4}{2\hbar^2} \tag{8.39}$$

S'il est bien entendu que ce calcul ne peut donner qu'un ordre de grandeur, la physique sous-jacente explique la raison profonde de la stabilité de l'atome : en raison des inégalités de Heisenberg, l'électron ne peut pas se trouver sur une orbite de rayon trop petit, sous peine d'acquérir une impulsion importante, qui fait croître son énergie cinétique. L'énergie de l'état fondamental est obtenue en recherchant le meilleur compromis possible entre énergie cinétique et énergie potentielle, de façon à obtenir le minimum de l'énergie totale.

8.1.5 Évolution du paquet d'ondes libre

Introduisons la dépendance du vecteur d'état par rapport au temps : le vecteur d'état est $|\varphi(0)\rangle \equiv |\varphi\rangle$ au temps $t = 0$ et $|\varphi(t)\rangle$ au temps t. La fonction d'onde $\varphi(x,t)$ au temps t est donc $\varphi(x,t) = \langle x|\varphi(t)\rangle$. Pour obtenir $|\varphi(t)\rangle$ en fonction de $|\varphi(0)\rangle$, nous avons besoin de l'équation d'évolution (4.11) et donc du hamiltonien H. Nous allons nous restreindre pour l'instant au cas où l'énergie potentielle est nulle, et où le hamiltonien se réduit à sa partie énergie cinétique K (7.66)

$$H = K = \frac{P^2}{2m} \tag{8.40}$$

Comme K et P commutent, les états propres de H peuvent être choisis parmi ceux de P

$$P|p\rangle = p|p\rangle \qquad H|p\rangle = \frac{P^2}{2m}|p\rangle = \frac{p^2}{2m}|p\rangle = E(p)|p\rangle \tag{8.41}$$

et par conséquent

$$\exp\Big[-\mathrm{i}\frac{Ht}{\hbar}\Big]|p\rangle = \exp\Big[-\mathrm{i}\frac{E(p)t}{\hbar}\Big]|p\rangle \tag{8.42}$$

Il est donc naturel d'exprimer $\langle x|\varphi(t)\rangle$ en fonction des composantes de $|\varphi(t)\rangle$ dans la base $|p\rangle$

$$\begin{aligned}\langle x|\varphi(t)\rangle &= \langle x|\exp\Big(-\mathrm{i}\frac{Ht}{\hbar}\Big)|\varphi(0)\rangle = \int \mathrm{d}p\,\langle x|p\rangle\langle p|\exp\Big(-\mathrm{i}\frac{Ht}{\hbar}\Big)|\varphi\rangle \\ &= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\mathrm{d}p\,\exp\Big(\mathrm{i}\,\frac{px}{\hbar} - \mathrm{i}\,\frac{E(p)t}{\hbar}\Big)\tilde{\varphi}(p)\end{aligned} \tag{8.43}$$

Afin d'éliminer les facteurs $\hbar$, on introduit le vecteur d'onde $k = p/\hbar$ et la fréquence $\omega(k)$

$$k = \frac{p}{\hbar} \qquad \omega(k) = \frac{E(\hbar k)}{\hbar} = \frac{\hbar k^2}{2m} \qquad A(k) = \sqrt{\hbar}\,\tilde{\varphi}(\hbar k)$$

pour écrire $\varphi(x,t)$ sous la forme

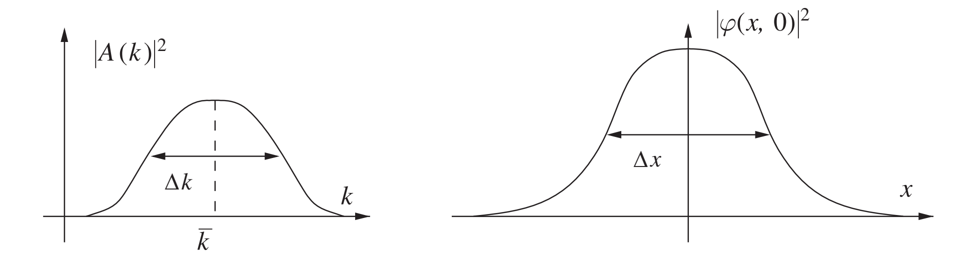


FIG. 8.2 – Dispersions du paquet d'ondes en k et en x.

$$\varphi(x,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}k A(k)\exp\Big(\mathrm{i}kx - \mathrm{i}\omega(k)t\Big) \tag{8.44}$$

L'allure qualitative de $|A(k)|^2$ et celle de $|\varphi(x,0)|^2$ sont représentées sur la figure 8.2. La fonction $|A(k)|^2$ est centrée à $k \simeq \overline{k}$ avec une largeur Δk. L'inégalité de Heisenberg (8.36) se traduit par

$$\Delta x\,\Delta k \geq \frac{1}{2} \tag{8.45}$$

Les cas limites sont :

- Particule de vecteur d'onde (ou d'impulsion) parfaitement défini = onde plane

$$A(k) = \delta(k - \overline{k}) \qquad \varphi(x,0) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{\mathrm{i}\overline{k}x} \tag{8.46}$$

- Particule localisée en $x = x_0$

$$A(k) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\mathrm{i}kx_0} \qquad\qquad \varphi(x,0) = \delta(x - x_0) \tag{8.47}$$

Rappelons que ni l'onde plane (8.46), ni l'état parfaitement localisé (8.47) ne correspondent à des états physiquement réalisables. Dans le cas (8.47) de la particule localisée, $|A(k)|^2$, qui est la probabilité d'observer une impulsion $\hbar k$, est indépendant de k, ce qui fait que la distribution de probabilité en p ne peut pas être normalisée. De même, dans le cas (8.46) de l'impulsion fixée, $|\varphi(x)|^2$ =cste et la probabilité de présence est uniforme sur l'axe des x : à nouveau la distribution de probabilité en x ne peut pas être normalisée. Pour un état physiquement réalisable, on doit avoir d'après (8.31)

$$\int_{-\infty}^{\infty} \mathrm{d}k\,|A(k)|^2 < \infty$$

Examinons maintenant l'évolution du paquet d'ondes au cours du temps. Pour évaluer (8.44), nous utilisons l'approximation de la phase stationnaire. Définissant $A(k) = |A(k)|\exp(\mathrm{i}\phi(k))$, la phase $\theta(k)$ de l'exponentielle dans (8.44) vaut

$$\theta(k) = kx - \omega(k)t + \phi(k)$$

Nous obtiendrons la contribution principale à l'intégrale (8.44) si la phase $\theta(k)$ est stationnaire dans la région $k \simeq \overline{k}$ où $|A(k)|$ est maximum : en effet, si $\theta(k)$ n'est pas stationnaire, l'exponentielle oscille rapidement et la contribution moyenne à l'intégrale (8.44) est nulle. Nous devons donc avoir

$$\left.\frac{\mathrm{d}\theta}{\mathrm{d}k}\right|_{k=\overline{k}} = x - t\left.\frac{\mathrm{d}\omega}{\mathrm{d}k}\right|_{k=\overline{k}} + \left.\frac{\mathrm{d}\phi}{\mathrm{d}k}\right|_{k=\overline{k}} = 0$$

Le centre du paquet d'ondes va se déplacer selon la loi

$$x = v_g(t - \tau) \tag{8.48}$$

où v_g est la vitesse de groupe, qui n'est autre que la vitesse moyenne $\overline{v}$ de la particule

$$v_g = \overline{v} = \left.\frac{\mathrm{d}\omega}{\mathrm{d}k}\right|_{k=\overline{k}} = \frac{\mathrm{d}}{\mathrm{d}k}\left.\frac{\hbar k^2}{2m}\right|_{k=\overline{k}} = \frac{\hbar\overline{k}}{m} = \frac{\overline{p}}{m} \tag{8.49}$$

Le temps τ qui détermine la position initiale $x_0 = -v_g\tau$ du centre du paquet d'ondes est

$$\tau = \frac{1}{v_g}\left.\frac{\mathrm{d}\phi}{\mathrm{d}k}\right|_{k=\overline{k}} = \hbar\left.\frac{\mathrm{d}\phi}{\mathrm{d}E}\right|_{k=\overline{k}} \tag{8.50}$$

Pour obtenir un résultat plus précis, nous pouvons récrire la phase en développant $\omega(k)$ au voisinage de $k = \overline{k}$

$$\begin{aligned}\theta(k) &= kx - \omega(\overline{k})t - (k - \overline{k})v_g t - \frac{1}{2}(k - \overline{k})^2\frac{\hbar}{m}\,t + \phi(k)\\ &= \omega(\overline{k})t + k(x - v_g t) - \frac{1}{2}(k - \overline{k})^2\frac{\hbar}{m}\,t + \phi(k)\end{aligned}$$

On obtient une forme très simple pour $\varphi(x,t)$ s'il est possible de négliger le terme quadratique en $(k-\overline{k})^2$

$$\begin{aligned} \varphi(x,t) &= \frac{1}{\sqrt{2\pi}} \exp[\mathrm{i}\omega(\overline{k})t] \int \mathrm{d}k A(k) \exp[\mathrm{i}k(x - v_g t)] \\ &= \exp[\mathrm{i}\omega(\overline{k})t]\varphi(x - v_g t, 0) \end{aligned} \tag{8.51}$$

Cette équation montre que, mis à part le facteur de phase $\exp[\mathrm{i}\omega(\overline{k})t]$, la fonction d'onde au temps t se déduit de celle au temps $t=0$ par la substitution $x \to x - v_g t$, c'est-à-dire que le paquet d'ondes se propage sans déformation dans la direction des x positifs – si $v_g > 0$ – à la vitesse v_g. Cependant, ce résultat est seulement approximatif puisque nous avons négligé le terme quadratique en $(k-\overline{k})^2$. Ce terme donne une contribution à la phase

$$-\frac{1}{2}(k-\overline{k})^2 \frac{\hbar}{m} t$$

qui doit rester $\ll 1$ dans le domaine où $|A(k)|$ est appréciable, si l'on veut se contenter de l'approximation linéaire. La contribution de ce terme pourra être négligée si

$$\frac{1}{2}(k-\overline{k})^2 \frac{\hbar t}{m} \ll 1$$

dans une région d'extension Δk autour de $\overline{k}$; pour que la déformation du paquet d'ondes soit faible, on doit avoir

$$t \ll \frac{2m}{\hbar(\Delta k)^2} = \frac{2m\hbar}{(\Delta p)^2} \tag{8.52}$$

Si cette condition n'est pas satisfaite, le paquet d'ondes se déforme en s'élargissant, tandis que son centre continue à se déplacer à la vitesse v_g : c'est le phénomène *d'étalement du paquet d'ondes*.

8.2 Équation de Schrödinger

8.2.1 Hamiltonien de l'équation de Schrödinger

Nous avons vu au § 7.4.1 que le hamiltonien indépendant du temps le plus général compatible avec l'invariance galiléenne en dimension $d=1$ était donné par (7.68)

$$H = \frac{P^2}{2m} + V(X) \tag{8.53}$$

où $K = P^2/(2m)$ est l'opérateur énergie cinétique et $V(X)$ l'opérateur énergie potentielle, ou plus brièvement le *potentiel*. Rappelons aussi l'équation d'évolution (4.11)

$$\mathrm{i}\hbar \frac{\mathrm{d}|\varphi(t)\rangle}{\mathrm{d}t} = H|\varphi(t)\rangle \tag{8.54}$$

Multiplions à gauche les deux membres de cette équation par le bra $\langle x|$ en utilisant (8.53) comme hamiltonien

$$\begin{aligned}
\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\langle x|\varphi(t)\rangle &= \mathrm{i}\hbar\frac{\partial}{\partial t}\varphi(x,t)\\
\langle x|P^2|\varphi(t)\rangle &= (P^2\varphi)(x,t) = \left(-\mathrm{i}\hbar\frac{\partial}{\partial x}\right)^2\varphi(x,t) = -\hbar^2\frac{\partial^2\varphi(x,t)}{\partial x^2}\\
\langle x|V(X)|\varphi(t)\rangle &= V(x)\varphi(x,t)
\end{aligned}$$

où nous avons utilisé (8.8) et (8.16). Nous obtenons donc l'équation de Schrödinger dépendant du temps

$$\boxed{\mathrm{i}\hbar\,\frac{\partial\varphi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\,\frac{\partial^2\varphi(x,t)}{\partial x^2} + V(x)\varphi(x,t)} \tag{8.55}$$

L'équation de Schrödinger est une équation d'onde pour la fonction d'onde $\varphi(x,t)$.

Comme le potentiel $V(X)$ est indépendant du temps, nous savons qu'il existe des solutions stationnaires de (8.54)

$$|\varphi(t)\rangle = \exp\Big(-\mathrm{i}\,\frac{Et}{\hbar}\Big)|\varphi(0)\rangle \qquad H|\varphi(0)\rangle = E|\varphi(0)\rangle \tag{8.56}$$

En multipliant à gauche par le bra $\langle x|$, l'équation $H|\varphi\rangle = E|\varphi\rangle$ devient l'équation de Schrödinger indépendante du temps

$$\boxed{\left[-\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial x^2} + V(x)\right]\varphi(x) = E\varphi(x)} \tag{8.57}$$

On peut généraliser (8.55) dans deux directions. Tout en restant compatible avec l'invariance galiléenne, il est possible de rajouter une dépendance temporelle au potentiel : $V(x) \to V(x,t)$. On peut aussi utiliser des potentiels dépendant de la vitesse, par exemple comme approximation d'effets relativistes. Dans ce cas, l'invariance galiléenne est perdue, et de plus il peut s'introduire des ambiguïtés lorsque l'on doit faire un choix pour l'ordre d'un produit d'opérateurs position et impulsion.

8.2.2 Probabilité de présence et vecteur courant

À la probabilité de présence $|\varphi(x,t)|^2$, on peut associer un vecteur courant $j(x,t)$, par analogie avec l'hydrodynamique ou l'électrodynamique. Rappelons l'exemple de l'hydrodynamique pour fixer les idées : soit[4] $\rho(\vec{r},t)$, la masse volumique d'un fluide compressible de masse totale M s'écoulant avec une vitesse locale $\vec{v}(\vec{r},t)$; le *courant* $\vec{\jmath}(\vec{r},t)$ est défini par

$$\vec{\jmath}(\vec{r},t) = \rho(\vec{r},t)\,\vec{v}(\vec{r},t) \tag{8.58}$$

4. Nous nous plaçons provisoirement en dimension $d = 3$.

Considérons une surface $\mathcal{S}$ limitant un volume $\mathcal{V}$, qui contient une masse $M(\mathcal{V})$ de fluide (figure 8.3). La masse $\mathrm{d}M(\mathcal{V})/\mathrm{d}t$ de fluide s'écoulant par unité de temps hors de $\mathcal{V}$ est égale au flux du courant à travers $\mathcal{S}$

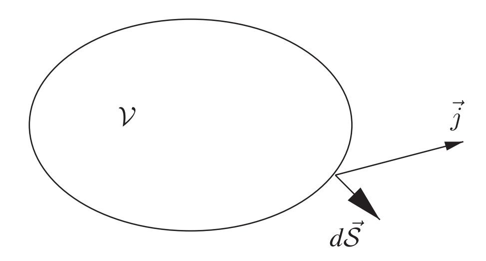


FIG. 8.3 – Courant et flux sortant d'un volume $\mathcal{V}$.

$$\frac{\mathrm{d}M(\mathcal{V})}{\mathrm{d}t} = \int_{\mathcal{S}} \vec{\jmath}\cdot \mathrm{d}\vec{\mathcal{S}} = \int_{\mathcal{V}} (\vec{\nabla}\cdot\vec{\jmath})\ \mathrm{d}^3 r$$

où nous avons utilisé le théorème de la divergence. Cette masse de fluide est aussi égale à moins la dérivée par rapport au temps de l'intégrale de la densité sur $\mathcal{V}$

$$\frac{\mathrm{d}M(\mathcal{V})}{\mathrm{d}t} = -\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{V}} \mathrm{d}^3 r\, \rho(\vec{r},t) = -\int_{\mathcal{V}} \mathrm{d}^3 r\, \frac{\partial \rho(\vec{r},t)}{\partial t}$$

Les deux expressions de $\mathrm{d}M(\mathcal{V})/\mathrm{d}t$ doivent être égales quel que soit le volume $\mathcal{V}$, ce qui implique que les intégrands doivent être égaux, d'où *l'équation de continuité*

$$\boxed{\frac{\partial \rho}{\partial t} + \vec{\nabla}\cdot\vec{\jmath} = 0} \tag{8.59}$$

et en revenant à la dimension $d = 1$

$$\frac{\partial \rho}{\partial t} + \frac{\partial j}{\partial x} = 0 \tag{8.60}$$

En électrodynamique, ρ est la densité de charge et $\vec{\jmath}$ la densité de courant, qui vérifient aussi une équation de continuité du type (8.59) traduisant la conservation locale de la charge électrique.

En mécanique quantique, on s'attend à trouver une équation de continuité du type (8.59), ou (8.60) à une dimension. En effet, si

$$\int_a^b \mathrm{d}x\, |\varphi(x,t)|^2$$

est la probabilité de trouver la particule au temps t dans l'intervalle $[a,b]$, cette probabilité va en général dépendre du temps. Si, par exemple, cette probabilité diminue, cela veut dire que la probabilité de trouver la particule

dans l'ensemble des deux intervalles complémentaires $]-\infty, a]$ et $[b, +\infty[$ doit augmenter, car quel que soit t, l'intégrale

$$\int_{-\infty}^{\infty} \mathrm{d}x\, |\varphi(x,t)|^2$$

est constante et égale à un. De la même façon, l'intégrale sur tout l'espace de la densité de fluide reste constante et égale à la masse totale M, ou, en électrodynamique, l'intégrale sur tout l'espace de la densité de charge reste constante et égale à la charge totale Q. L'analogue de la densité en mécanique quantique est $\rho(x,t) = |\varphi(x,t)|^2$; cependant, il s'agit d'une densité de probabilité, et non d'une véritable densité. Nous recherchons un courant $j(x,t)$ vérifiant (8.60), qui sera aussi un courant de probabilité et non un courant véritable. La forme de ce courant est suggérée par le raisonnement suivant : en hydrodynamique, la vitesse moyenne $\langle v(t)\rangle$ du fluide (ou vitesse du centre de masse) est donnée par

$$\langle v(t)\rangle = \frac{1}{M}\int \rho(x,t)v(x,t)\mathrm{d}x = \frac{1}{M}\int j(x,t)\mathrm{d}x \tag{8.61}$$

En mécanique quantique, l'opérateur vitesse est d'après (7.61)

$$\dot{X} = \frac{\mathrm{i}}{\hbar}\,[H, X] = \frac{P}{m}$$

et sa valeur moyenne est

$$\langle \dot{X}\rangle(t) = \langle\varphi(t)\Big|\frac{P}{m}\Big|\varphi(t)\rangle = \int \mathrm{d}x\, \varphi^*(x,t)\frac{\hbar}{\mathrm{i}m}\,\frac{\partial\varphi(x,t)}{\partial x}$$

où nous avons utilisé (8.9) et (8.16). L'intégrand dans cette équation est en général complexe et ne convient pas pour le courant. Une intégration par parties permet de construire un courant réel

$$\langle \dot{X}\rangle(t) = \frac{\hbar}{2\mathrm{i}m}\int \mathrm{d}x\, \left(\varphi^*(x,t)\frac{\partial\varphi(x,t)}{\partial x} - \varphi(x,t)\frac{\partial\varphi^*(x,t)}{\partial x}\right) \tag{8.62}$$

La comparaison de (8.61) pour $M = 1$ avec (8.62) suggère la forme suivante pour le courant $j(x,t)$

$$\boxed{j = \frac{\hbar}{2\mathrm{i}m}\left(\varphi^*(x,t)\frac{\partial\varphi(x,t)}{\partial x} - \varphi(x,t)\frac{\partial\varphi^*(x,t)}{\partial x}\right) = \mathrm{Re}\,\left(\frac{\hbar}{\mathrm{i}m}\,\varphi^*(x,t)\frac{\partial\varphi(x,t)}{\partial x}\right)} \tag{8.63}$$

Afin de nous familiariser avec cette expression peu intuitive, examinons le cas d'une onde plane

$$\varphi(x) = A\,\mathrm{e}^{\mathrm{i}px/\hbar}$$

La densité est $\rho(x) = |A|^2$. Le courant vaut

$$j(x) = \mathrm{Re}\left(\frac{\hbar}{\mathrm{i}m} A^* \, \mathrm{e}^{-\mathrm{i}px/\hbar} \left[\frac{\mathrm{i}p}{\hbar}\right] A \, \mathrm{e}^{\mathrm{i}px/\hbar}\right) = |A|^2 \, \frac{p}{m} \tag{8.64}$$

et s'interprète donc comme : courant = densité×vitesse. Le courant est dirigé vers la droite si $p > 0$ et vers la gauche si $p < 0$. Lorsque la fonction d'onde est indépendante du temps, comme dans le cas de l'onde plane, le courant est nécessairement indépendant de x puisque $\partial\rho/\partial t = 0 \Longrightarrow \partial j/\partial x = 0$. Il reste à s'assurer que le courant (8.63) vérifie bien l'équation de continuité (8.60). D'une part,

$$\frac{\partial j}{\partial x} = \frac{\hbar}{2\mathrm{i}m}\left[\varphi^* \frac{\partial^2 \varphi}{\partial x^2} - \varphi \frac{\partial^2 \varphi^*}{\partial x^2}\right] = \frac{\mathrm{i}}{\hbar}\left[\varphi^*(H\varphi) - \varphi(H\varphi)^*\right]$$

où nous avons utilisé

$$\frac{\hbar}{2\mathrm{i}m}\frac{\partial^2 \varphi}{\partial x^2} = \frac{\mathrm{i}}{\hbar}\left[(H - V)\varphi\right]$$

et le fait que V est une fonction réelle de x et t. D'autre part,

$$\frac{\partial}{\partial t}|\varphi(x,t)|^2 = \varphi^* \frac{\partial \varphi}{\partial t} + \varphi \frac{\partial \varphi^*}{\partial t} = \frac{1}{\mathrm{i}\hbar}\left[\varphi^*(H\varphi) - \varphi(H\varphi)^*\right]$$

ce qui montre que

$$\frac{\partial}{\partial t}|\varphi(x,t)|^2 + \frac{\partial}{\partial x} j(x,t) = 0 \tag{8.65}$$

8.3 Résolution de l'équation de Schrödinger indépendante du temps

8.3.1 Généralités

Nous nous proposons de chercher les solutions de l'équation de Schrödinger indépendante du temps (8.57), c'est-à dire déterminer les valeurs propres E et les fonctions propres correspondantes $\varphi(x)$. Commençons par le cas le plus simple où le potentiel $V(x) = 0$. L'équation (8.57) devient

$$\left(\frac{\partial^2}{\partial x^2} + \frac{2mE}{\hbar^2}\right)\varphi(x) = 0 \tag{8.66}$$

La solution générale de cette équation est bien sûr une combinaison d'ondes planes, avec $p = \sqrt{2mE} > 0$

$$\varphi(x) = A \, \mathrm{e}^{\mathrm{i}px/\hbar} + B \, \mathrm{e}^{-\mathrm{i}px/\hbar} \tag{8.67}$$

se propageant dans la direction des x positifs avec une amplitude A et des x négatifs avec une amplitude B. Bien que la solution (8.67) soit indépendante

du temps, elle génère un courant stationnaire[5] composé d'après (8.64) d'un terme $|A|^2 p/m$ dirigé vers les x positifs et d'un terme $-|B|^2 p/m$ dirigé vers les x négatifs. Aux solutions indépendantes du temps $\exp(\pm \mathrm{i} px/\hbar)$ correspondent des solutions de (8.55) dépendant du temps $\exp[\mathrm{i}(\pm px - E(p)t)/\hbar]$ qui sont des ondes progressives se propageant dans les directions des x positifs ou des x négatifs. Avec les ondes progressives $\exp[\mathrm{i}(+px - E(p)t)/\hbar]$, on forme des paquets d'ondes se propageant dans la direction des x positifs : on dira que ces paquets d'ondes proviennent d'une *source* de particules à $x = -\infty$. Avec les ondes progressives $\exp[\mathrm{i}(-px - E(p)t)/\hbar]$, on forme des paquets d'ondes se propageant dans la direction des x négatifs : on dira qu'il y a une source de particules à $x = +\infty$.

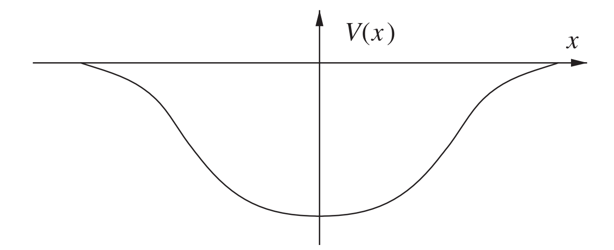


FIG. 8.4 – Puits de potentiel.

Passons au cas où $V(x) \neq 0$ et, pour fixer les idées, supposons que $V(x)$ a la forme de la figure 8.4 : $V(x)$ est un "puits de potentiel" et $V(x) \to 0$ si $x \to \pm\infty$. En mécanique classique, suivant la discussion du § 1.5.1, ce potentiel a des états liés si $E < 0$ et des états de diffusion si $E > 0$. Pour $E < 0$, la particule classique reste confinée sur un intervalle fini de l'axe des x, pour $E > 0$, elle part à l'infini. La région de l'axe des x qui est permise à la particule classique est celle où $E > V(x)$, et où son impulsion $p(x)$ est réelle

$$p(x) = \pm\sqrt{2m(E - V(x))} \tag{8.68}$$

tandis que la région $E < V(x)$, où son impulsion est imaginaire

$$p(x) = \pm \mathrm{i}\sqrt{2m(V(x) - E)} \tag{8.69}$$

lui est interdite. Nous allons voir que ce comportement classique se reflète dans le comportement quantique : la forme des solutions de (8.57) sera différente selon que $p(x)$ est réel ou imaginaire. Pour que $\varphi(x)$ soit une solution acceptable, il ne suffit pas que $\varphi(x)$ vérifie formellement (8.57), il faut aussi que $\varphi(x)$ soit normalisable

$$\int_{-\infty}^{\infty} \mathrm{d}x\, |\varphi(x)|^2 < \infty$$

5. Un exemple de courant stationnaire est le courant continu en électricité.

C'est cette condition que nous allons utiliser pour obtenir les états liés. Pour les états de diffusion, cette condition est trop forte : nous avons vu que pour $V(x) = 0$, les solutions de (8.57) sont des ondes planes, non normalisables. Pour $x \to \pm\infty$, nous nous attendons à un comportement d'onde plane pour les solutions de (8.57), puisque le potentiel s'annule à l'infini. Pour les états de diffusion $E > 0$ du potentiel de la figure 8.4, nous nous contenterons d'exiger un comportement d'onde plane à l'infini : il ne faut pas exiger davantage de la solution en présence de potentiel que de la solution en l'absence de potentiel !

8.3.2 Réflexion et transmission par une marche de potentiel

Dans la suite de cette section, nous allons nous intéresser au cas où le potentiel est constant par morceaux : $V(x)$ est constant dans un intervalle et saute brusquement à une autre constante en certains points. Ce type de potentiel représente une bonne approximation d'un potentiel réaliste dans certains cas, et il peut être utilisé pour approcher un potentiel variant continûment dans d'autres cas. Comme le potentiel présente des discontinuités, il est nécessaire d'examiner le comportement de la fonction d'onde au voisinage d'une telle discontinuité. Montrons que la fonction d'onde $\varphi(x)$ et sa dérivée $\varphi'(x)$ sont continues si le potentiel présente une discontinuité *finie* V_0 au point $x = x_0$ (figure 8.5). Comme $|\varphi(x)|^2$ doit être intégrable en x_0, $|\varphi(x)|$ l'est également. Il sera commode de récrire l'équation de Schrödinger indépendante du temps (8.57) sous la forme

$$\left(\frac{\partial^2}{\partial x^2} + \frac{2m(E - V(x))}{\hbar^2}\right)\varphi(x) = 0 \tag{8.70}$$

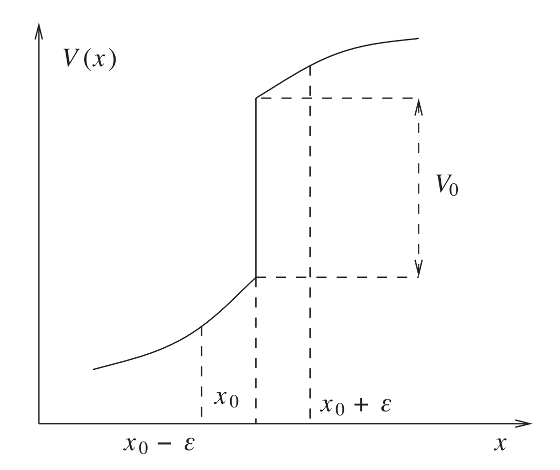


Fig. 8.5 – Discontinuité de potentiel.

Au voisinage de la discontinuité, nous pouvons déduire le comportement de $\varphi'(x)$ grâce à

$$\varphi'(x_0+\varepsilon)-\varphi'(x_0-\varepsilon)=\int_{x_0-\varepsilon}^{x_0+\varepsilon}\frac{\partial^2\varphi(x)}{\partial x^2}\,\mathrm{d}x=\int_{x_0-\varepsilon}^{x_0+\varepsilon}\left[\frac{2m(V(x)-E)}{\hbar^2}\right]\varphi(x)$$

La seconde intégrale est bien définie car $\varphi(x)$ est intégrable ; cette intégrale doit donc tendre vers zéro avec ε, ce qui montre que $\varphi'(x)$ et *a fortiori* $\varphi(x)$ sont continues tant que la discontinuité V_0 reste finie.

Au lieu d'écrire les conditions de continuité de $\varphi(x)$ et de $\varphi'(x)$, il est souvent commode d'écrire celles de $\varphi(x)$ et de sa dérivée logarithmique $\varphi'(x)/\varphi(x)$. Une conséquence immédiate de ces conditions est que le courant $j(x)$ est égal à la même constante de part et d'autre de x_0, ce que l'on voit à partir de l'expression suivante de $j(x)$

$$j(x)=\frac{\hbar}{2\mathrm{i}m}|\varphi|^2\left[\frac{\varphi'}{\varphi}-\left(\frac{\varphi'}{\varphi}\right)^*\right]$$

Comme application de ces conditions de continuité, prenons le cas d'une “marche de potentiel” (figure 8.6)

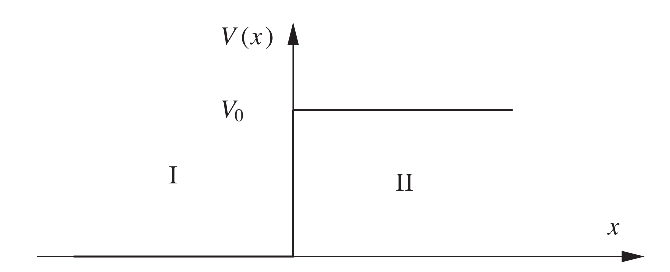


FIG. 8.6 – Marche de potentiel.

$$\begin{array}{lcclll} \text{région I} & V(x) & = & 0 & \text{pour} & x<0 \\ \text{région II} & V(x) & = & V_0 & \text{pour} & x>0 \end{array}$$

Pour fixer les idées, on choisit $0<E<V_0$. Si l'on définit k et κ par

$$k=\sqrt{\frac{2mE}{\hbar^2}}\qquad \kappa=\sqrt{\frac{2m(V_0-E)}{\hbar^2}} \tag{8.71}$$

les solutions de (8.70) s'écrivent dans les régions I et II

$$\text{I}\qquad \varphi(x)=A\,\mathrm{e}^{\mathrm{i}kx}+B\,\mathrm{e}^{-\mathrm{i}kx} \tag{8.72}$$

$$\text{II}\qquad \varphi(x)=C\,\mathrm{e}^{-\kappa x}+D\,\mathrm{e}^{\kappa x} \tag{8.73}$$

Si $V(x)$ reste égal à V_0 pour tout $x > 0$, le comportement (8.73) de la fonction d'onde reste valable quel que soit $x > 0$. Il est alors nécessaire que $D = 0$, car dans le cas contraire, la fonction $|\varphi(x)|^2$ se comporterait comme $\exp(2\kappa x)$ pour $x \to \infty$: un comportement constant en module comme celui d'une onde plane est acceptable, mais pas un comportement aussi divergent. Dans ces conditions, la continuité à $x = 0$ de φ et celle de sa dérivée logarithmique s'écrivent

$$\varphi: \; C = A + B \qquad \frac{\varphi'}{\varphi}: \; -\kappa = \frac{\mathrm{i}k(A - B)}{A + B}$$

Les coefficients A et B sont *a priori* définis à une constante multiplicative près puisque l'on n'a fait aucune hypothèse sur la région $x > 0$. On peut fixer arbitrairement $A = 1$, et la solution pour les deux autres coefficients est alors

$$B = -\frac{\kappa + \mathrm{i}k}{\kappa - \mathrm{i}k} \qquad C = -\frac{2\mathrm{i}k}{\kappa - \mathrm{i}k} \tag{8.74}$$

Nous pouvons déduire de ces expressions le cas limite où $V_0 \to \infty$, qui correspond à une barrière infranchissable par une particule classique quelle que soit son énergie, ou barrière de potentiel infinie. L'équation (8.71) montre alors que $\kappa \to \infty$, et (8.74) que $B \to -1$ et $C \to 0$. La fonction d'onde s'annule dans la région II, en restant continue au point $x = 0$. En revanche, sa dérivée $\varphi'(x)$ est discontinue en ce point.

8.3.3 États liés du puits carré

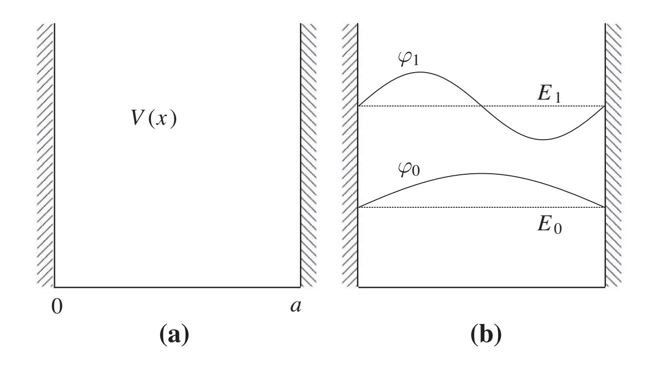


FIG. 8.7 – Puits carré infini et fonctions d'onde de ses deux premiers niveaux.

Comme premier exemple d'états liés, étudions ceux du puits carré infini (figure 8.7)

$$\begin{aligned} V(x) &= 0 && 0 \le x \le a \\ V(x) &= +\infty && x < 0 \text{ ou } x > a \end{aligned}$$

Les barrières de potentiel en $x = 0$ et $x = a$ sont infinies : une particule classique est enfermée dans la région $0 \le x \le a$, quelle que soit son énergie. D'après la discussion précédente, la fonction d'onde de la particule quantique s'annule en dehors de l'intervalle $[0, a]$ et une particule quantique est également strictement confinée dans $[0, a]$: sa probabilité de présence est nulle en dehors de l'intervalle $[0, a]$. En raison de l'annulation de la fonction d'onde à $x = 0$, les solutions de (8.70) sont de la forme

$$\varphi(x) = A \sin(kx) \qquad k = \sqrt{\frac{2mE}{\hbar^2}}$$

et comme $\varphi(x)$ doit aussi s'annnuler à $x = a$, les valeurs de k sont de la forme

$$k = k_n = \frac{\pi(n+1)}{a} \qquad n = 0, 1, 2, 3, \ \ldots \tag{8.75}$$

Nous voyons que l'énergie prend des valeurs discrètes étiquetées par un nombre entier positif[6] n

$$\boxed{E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2}{2m}\left(\frac{\pi}{a}\right)^2 (n+1)^2 \qquad n = 0, 1, 2, 3, \ \ldots} \tag{8.76}$$

En d'autres termes, nous venons de montrer que les niveaux d'énergie du puits infini sont *quantifiés*, et c'est le premier exemple où nous démontrons explicitement cette quantification. La fonction d'onde correctement normalisée correspondant au niveau E_n est

$$\varphi_n(x) = \sqrt{\frac{2}{a}}\, \sin \frac{\pi(n+1)x}{a} \tag{8.77}$$

Il est facile de vérifier que deux fonctions d'onde $\varphi_n(x)$ et $\varphi_m(x)$ sont orthogonales pour $n \neq m$. Les valeurs k_n et $-k_n$ correspondent au même état physique car la substitution $k_n \to -k_n$ conduit à un simple changement de signe de la fonction d'onde, qui est un facteur de phase : c'est pourquoi nous n'avons pas pris en compte les valeurs négatives de n dans (8.75). Remarquons également que la fonction d'onde $\varphi_n(x)$ s'annule n fois dans l'intervalle $[0, a]$: on dit que la fonction d'onde a n *nœuds* dans cet intervalle. Le nombre de nœuds donne une classification des niveaux par valeurs croissantes de l'énergie : plus l'énergie est grande et plus la fonction d'onde a de nœuds. Ceci est un résultat général que nous admettrons lorsque le potentiel $V(x)$ est suffisamment régulier, ce que nous supposerons toujours : si E_n est l'énergie du niveau n, la fonction d'onde correspondante aura n noeuds. La fonction d'onde de l'état fondamental E_0 ne s'annule pas. Une autre remarque est que

6. Notre convention (qui n'est pas la convention habituelle dans ce cas précis) est choisie de telle sorte $n = 0$ corresponde à l'état fondamental, ce qui permet de se conformer à une convention usuelle : en général, l'énergie de l'état fondamental est notée E_0.

l'inégalité de Heisenberg permet de trouver l'ordre de grandeur de l'énergie de l'état fondamental. En effet, $p \sim \hbar/x \sim \hbar/a$, ce qui donne

$$E = \frac{p^2}{2m} \sim \frac{\hbar^2}{2ma^2}$$

en accord avec (8.76) pour $n = 0$ à un facteur π^2 près. Contrairement au cas de l'atome d'hydrogène, le résultat heuristique diffère du résultat exact par un facteur ~ 10 : ceci provient de la variation brutale du potentiel à $x = 0$ et $x = a$ qui oblige la fonction d'onde à s'annuler de façon brusque, avec comme conséquence une énergie cinétique importante. En effet, l'énergie cinétique moyenne $\langle K \rangle_\varphi$ dans l'état φ est

$$\langle K \rangle_\varphi = \langle \varphi | K | \varphi \rangle = -\frac{\hbar^2}{2m} \int \mathrm{d}x \, \varphi^*(x) \frac{\mathrm{d}^2\varphi(x)}{\mathrm{d}x^2}$$

et elle est d'autant plus importante que la dérivée seconde de $\varphi(x)$ est grande.

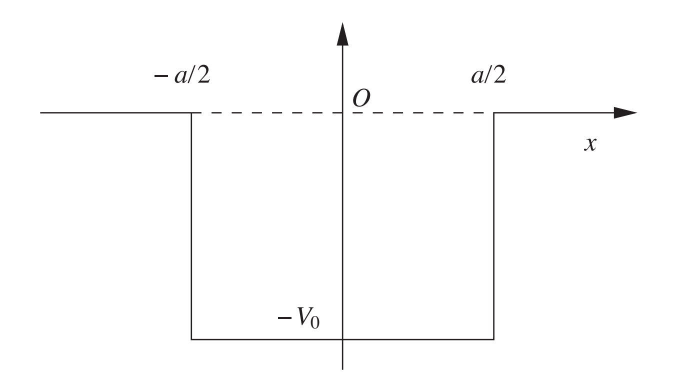


FIG. 8.8 – Puits carré fini.

Déterminons maintenant les niveaux d'énergie du puits carré fini (figure 8.8)

$$\begin{aligned} V(x) &= 0 \qquad & |x| > a/2 \\ V(x) &= -V_0 \qquad & |x| < a/2 \end{aligned}$$

On cherche les états liés, et on doit donc choisir l'énergie dans l'intervalle $[-V_0, 0]$. Définissons k et κ par

$$\kappa = \sqrt{-\frac{2mE}{\hbar^2}} \qquad k = \sqrt{\frac{2m(V_0 + E)}{\hbar^2}} \qquad 0 \le \kappa^2 \le \frac{2mV_0}{\hbar^2} \tag{8.78}$$

Le potentiel $V(x)$ est invariant dans l'opération parité $\Pi : x \to -x$. En effet, $V(-x) = V(x)$, ce qui entraîne l'invariance du hamiltonien dans cette

opération : $H(-x) = H(x)$. D'après la discussion du § 7.3.3, on peut chercher les vecteurs propres $|\varphi_\pm\rangle$ de H qui sont pairs ou impairs dans l'opération parité

$$\Pi|\varphi_\pm\rangle = \pm|\varphi_\pm\rangle$$

En termes de fonction d'onde, si $\langle x|\varphi_\pm\rangle = \varphi_\pm(x)$

$$\varphi_+(-x) = \varphi(x) \qquad \varphi_-(-x) = -\varphi_-(x)$$

En effet, compte tenu de $\Pi|x\rangle = |-x\rangle$

$$\begin{aligned}\langle x|\Pi|\varphi_\pm\rangle &= \langle -x|\varphi_\pm\rangle = \varphi_\pm(-x) \\ &= \pm\langle x|\varphi_\pm\rangle = \pm\varphi_\pm(x)\end{aligned}$$

Les solutions de l'équation de Schrödinger (8.70) se divisent donc en solutions paires et impaires. Écrivons les fonctions d'ondes paires et impaires dans les régions I : $x < -a/2$, II : $|x| < a/2$ et III : $x > a/2$; la colonne du milieu du tableau ci-dessous donne les fonctions d'onde paires, la colonne de droite les fonctions d'onde impaires

I	$A\,\mathrm{e}^{-\kappa\|x\|}$	$-\,A'\mathrm{e}^{-\kappa\|x\|}$
II	$B\cos(kx)$	$B'\sin(kx)$
III	$A\,\mathrm{e}^{-\kappa x}$	$A'\mathrm{e}^{-\kappa x}$

Les conditions de continuité de φ'/φ au point $x = a/2$ donnent

$$\text{Solutions paires} \qquad \kappa = k\tan(ka/2) \tag{8.79}$$

$$\text{Solutions impaires} \qquad \kappa = -k\cot(ka/2) \tag{8.80}$$

La solution graphique de ces équations est représentée sur la figure 8.9. On voit qu'il existe un nombre fini d'états liés, et qu'il existe toujours au moins un état lié.

8.3.4 Diffusion par un potentiel

Après l'étude des états liés, nous passons à celle des états de diffusion. Nous nous proposons d'étudier le comportement d'une particule quand elle traverse une barrière de potentiel de la forme suivante (figure 8.10).

Région I	$x < -a$	:	$V(x) = 0$
Région II	$-\,b < x < b$	:	$V(x) = V_0$
Région III	$x > a$	:	$V(x) = 0$

Nous nous limitons au cas $E > V_0$, le cas $E < V_0$ sera examiné au § 12.4.5. Définissant

$$k = \sqrt{\frac{2mE}{\hbar^2}} \qquad k' = \sqrt{\frac{2m(E-V_0)}{\hbar^2}} \tag{8.81}$$

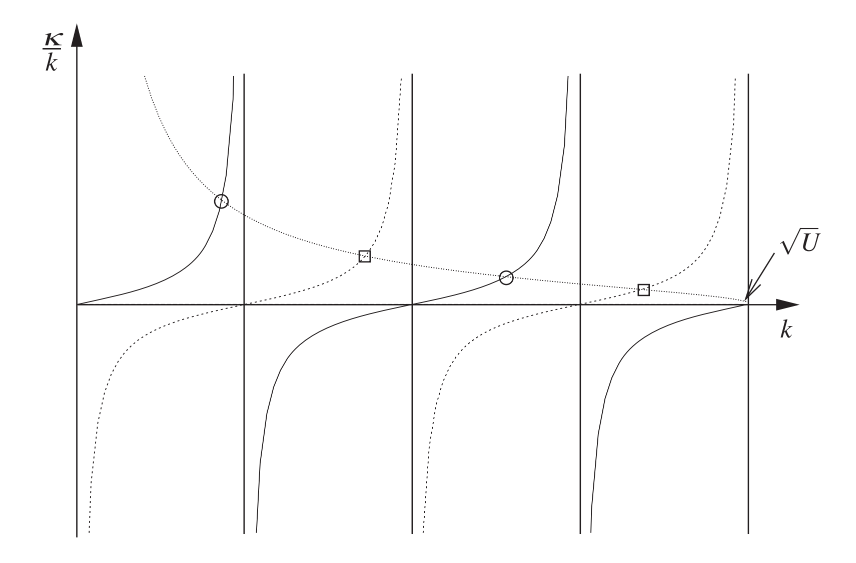


Fig. 8.9 – Détermination graphique des états liés du puits carré fini par intersection des courbes $\tan ka/2$: solutions paires (8.79) et $-\cot ka/2$: solutions impaires (8.80) avec $\sqrt{U-k^2}/k$, où $U = 2mV_0/\hbar^2$.

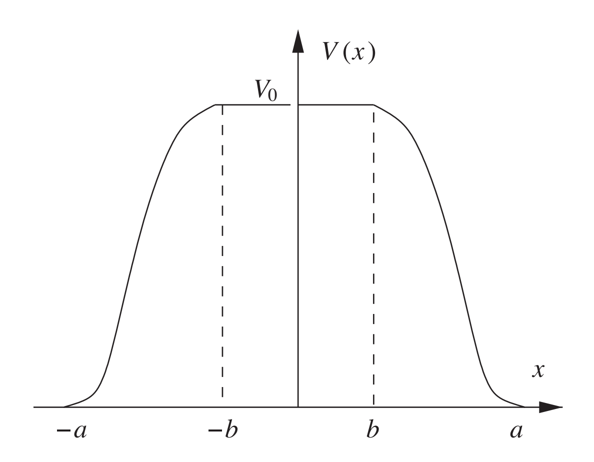


Fig. 8.10 – Une barrière de potentiel.

les fonctions d'onde dans les trois régions sont

$$\text{I} : \; x < -a \qquad \varphi(x) = A\,\mathrm{e}^{\mathrm{i}kx} + B\,\mathrm{e}^{-\mathrm{i}kx} \tag{8.82}$$

$$\text{II} : \; -b \leq x \leq b \qquad \varphi(x) = F\,\mathrm{e}^{\mathrm{i}k'x} + G\,\mathrm{e}^{-\mathrm{i}k'x} \tag{8.83}$$

$$\text{III} : \; x > a \qquad \varphi(x) = C\,\mathrm{e}^{\mathrm{i}kx} + D\,\mathrm{e}^{-\mathrm{i}kx} \tag{8.84}$$

Nous examinons d'abord le passage de la région I à la région II, c'est-à-dire l'intervalle $[-a, -b]$. Comme l'équation de Schrödinger est linéaire, A et B sont reliés linéairement à F et G, ce que l'on peut écrire sous forme matricielle

$$\begin{pmatrix} A \\ B \end{pmatrix} = R \begin{pmatrix} F \\ G \end{pmatrix} \tag{8.85}$$

où R est une matrice 2×2. Il est possible de déterminer des propriétés de R sans résoudre explicitement l'équation de Schrödinger. On observe en effet que si $\varphi(x)$ est solution de l'équation de Schrödinger indépendante du temps (8.57), alors son complexe conjugué $\varphi^*(x)$ est également solution de cette équation parce que le potentiel $V(x)$ est réel. Cette propriété est une conséquence de l'invariance par renversement du sens du temps comme on le montre dans l'appendice A2. La fonction $\varphi^*(x)$ dans les régions I et II est

$$\text{I} \quad \varphi^*(x) = A^* \mathrm{e}^{-\mathrm{i}kx} + B^* \mathrm{e}^{\mathrm{i}kx} \tag{8.86}$$

$$\text{II} \quad \varphi^*(x) = F^* \mathrm{e}^{-\mathrm{i}k'x} + G^* \mathrm{e}^{\mathrm{i}k'x} \tag{8.87}$$

Comparant les coefficients de $\exp(\pm \mathrm{i}kx)$ et $\exp(\pm \mathrm{i}k'x)$ avec ceux de (8.82) et (8.83), on déduit de (8.85)

$$\begin{pmatrix} B^* \\ A^* \end{pmatrix} = R \begin{pmatrix} G^* \\ F^* \end{pmatrix}$$

soit

$$R^*_{11} = R_{22} \qquad R^*_{12} = R_{21}$$

On peut donc écrire la matrice R en fonction de deux nombres complexes α et β

$$R = \sqrt{\frac{k'}{k}} \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \tag{8.88}$$

La raison pour l'introduction du facteur *a priori* arbitraire $\sqrt{k'/k}$ va apparaître dans un instant. La conservation du courant dans les régions I et II se traduit par la relation

$$k(|A|^2 - |B|^2) = k'(|F|^2 - |G|^2)$$

Calculons le courant dans la région I en exprimant A et B fonction de F et G

$$\begin{aligned} k(|A|^2 - |B|^2) &= k\,\frac{k'}{k}\left(|\alpha F + \beta G|^2 - |\beta^* F + \alpha^* G|^2\right) \\ &= k'\left(|\alpha|^2 - |\beta|^2\right)\left(|F|^2 - |G|^2\right) \end{aligned}$$

ce qui implique $|\alpha|^2 - |\beta|^2 = 1$: la matrice $\sqrt{k/k'}\,R$ est de déterminant un. On voit l'intérêt du coefficient $\sqrt{k'/k}$ dans (8.88) : en raison de la variation de la

vitesse entre les régions I et II, c'est la matrice $\sqrt{k/k'}\,R$ qui a les propriétés les plus simples.

Pour passer de la région II à la région III, on remarque que le passage de la région III vers la région II est identique au passage de I à II, à condition de changer $x \to -x$ et $k \to k'$, et donc

$$\begin{pmatrix} D \\ C \end{pmatrix} = R \begin{pmatrix} G \\ F \end{pmatrix}$$

soit

$$\begin{pmatrix} F \\ G \end{pmatrix} = \sqrt{\frac{k}{k'}} \begin{pmatrix} \alpha & -\beta^* \\ -\beta & \alpha^* \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \tilde{R} \begin{pmatrix} C \\ D \end{pmatrix} \tag{8.89}$$

Combinant (8.88) et (8.89), on obtient le passage $(A, B) \to (C, D)$

$$\begin{pmatrix} A \\ B \end{pmatrix} = R \begin{pmatrix} F \\ G \end{pmatrix} = R\tilde{R} \begin{pmatrix} C \\ D \end{pmatrix} = M \begin{pmatrix} C \\ D \end{pmatrix} \tag{8.90}$$

et l'on a donc $M = R\tilde{R}$, où M est la *matrice de passage*. En utilisant les formes trouvées pour R et $\tilde{R}$, on déduit

$$M_{11} = M_{22}^* \qquad M_{12} = M_{21}^* \qquad \det M = 1 \tag{8.91}$$

À titre d'exercice, le lecteur pourra démontrer directement ces relations à partir de l'invariance par renversement du sens du temps et la conservation du courant, sans passer par l'intermédiaire des matrices R et $\tilde{R}$. La forme générale de M est par conséquent

$$M = \begin{pmatrix} \gamma & \delta \\ \delta^* & \gamma^* \end{pmatrix} \qquad |\gamma|^2 - |\delta|^2 = 1 \tag{8.92}$$

Cette expression de M est indépendante de la forme du potentiel, pourvu que celui-ci s'annule suffisamment rapidement pour $x \to \pm\infty$.

Il reste une propriété générale de M que nous n'avons pas encore exploitée : lorsque le potentiel est invariant par parité, $V(x) = V(-x)$, l'opération parité $x \to -x$ échange les régions I et III. Si $\varphi(x)$ est la solution initiale et $\chi(x) = \varphi(-x)$, nous avons

$$\begin{array}{rrcl} \text{région I} & \chi(x) & = & C\,\mathrm{e}^{-\mathrm{i}kx} + D\,\mathrm{e}^{\mathrm{i}kx} \\ \text{région III} & \chi(x) & = & A\,\mathrm{e}^{-\mathrm{i}kx} + B\,\mathrm{e}^{\mathrm{i}kx} \end{array}$$

et la relation entre les différents coefficients est maintenant

$$\begin{pmatrix} C \\ D \end{pmatrix} = M \begin{pmatrix} B \\ A \end{pmatrix}$$

ou encore

$$\begin{pmatrix} B \\ A \end{pmatrix} = M^{-1} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} M_{22} & -M_{12} \\ -M_{21} & M_{11} \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix}$$

Nous avons utilisé $\det M = 1$. Comparant avec (8.91), nous en déduisons que M est une matrice antisymétrique : $M_{12} = -M_{21}$, ce qui, joint à $M_{12}^* = M_{21}$, implique que δ est imaginaire pur, $\delta = \mathrm{i}\eta$, η réel.

Au chapitre 13, nous étudierons la théorie de la diffusion dans l'espace à trois dimensions. Nous verrons qu'un outil important de cette théorie est la matrice S, que nous allons introduire dans le cas plus simple à une dimension. Nous supposons le potentiel de forme quelconque s'annulant[7] dans la région $|x| > a$. Des sources de particules à $x = -\infty$ et $x = +\infty$ génèrent des ondes planes $\exp(\mathrm{i}kx)$ et $\exp(-\mathrm{i}kx)$ dans les régions $x < -a$ et $x > a$, respectivement : nous appellerons ces ondes les *ondes entrantes*. Ces ondes entrantes peuvent être réfléchies ou transmises et donner des ondes sortantes $\exp(-\mathrm{i}kx)$ dans la région $x < -a$ et $\exp(\mathrm{i}kx)$ dans la région $x > a$. Par définition, la matrice S relie les coefficients A et D des ondes entrantes (se propageant *vers* la barrière) aux coefficients B et C des ondes sortantes (*cf.* (8.82) et (8.84))

$$\begin{pmatrix} B \\ C \end{pmatrix} = S \begin{pmatrix} A \\ D \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A \\ D \end{pmatrix} \tag{8.93}$$

La matrice S peut, bien sûr, s'exprimer en fonction de M. Cependant, il est plus instructif de répéter les raisonnements qui nous ont donné les propriétés générales de M.

(i) Conservation du courant

$$|A|^2 - |B|^2 = |C|^2 - |D|^2 \Longrightarrow |A|^2 + |D|^2 = |B^2| + |C|^2$$

Cette équation montre que S conserve la norme, et S est donc unitaire[8].

(ii) $\varphi^*(x)$ solution de l'équation de Schrödinger

$$\begin{pmatrix} A^* \\ D^* \end{pmatrix} = S \begin{pmatrix} B^* \\ A^* \end{pmatrix} \Longrightarrow \begin{pmatrix} B \\ A \end{pmatrix} = (S^*)^{-1} \begin{pmatrix} A \\ D \end{pmatrix}$$

d'où

$$S = (S^*)^{-1} = (S^{-1})^* = (S^\dagger)^* = S^T \tag{8.94}$$

La matrice S est symétrique : $S_{12} = S_{21}$. L'opération de conjugaison complexe échange les ondes entrantes et les ondes sortantes, ce qui correspond à renverser le sens du temps : la propriété de symétrie $S_{12} = S_{21}$ est donc liée à l'invariance par renversement du sens du temps.

Si on le souhaite, il est immédiat d'exprimer la matrice S en fonction de la matrice M.

Un aspect intéressant de la matrice S est qu'elle permet de relier la diffusion aux états liés, et plus généralement aux résonances (exercice 13.6.4).

7. On peut généraliser au cas d'un potentiel s'annulant suffisamment vite pour $x \to \pm\infty$.

8. Le raisonnement ne vaut que pour la dimension finie : nous avons seulement prouvé que S est une isométrie, ce qui suffit à en faire un opérateur unitaire en dimension finie. Il se trouve que S est unitaire même en dimension infinie, mais il faut un argument supplémentaire pour le prouver (section 13.5).

Prenant un puits de potentiel de forme quelconque (mais tel que $V(x) = 0$ en dehors d'un intervalle fini pour simplifier la discussion), nous choisissons $E < 0$ avec $\kappa = -\mathrm{i}k$ donné par (8.81). Les fonctions d'onde dans les régions I et III sont

$$\begin{array}{lrcl} \text{Région I} & \varphi(x) & = & A\,\mathrm{e}^{-\kappa x} + B\,\mathrm{e}^{\kappa x} \\ \text{Région III} & \varphi(x) & = & C\,\mathrm{e}^{-\kappa x} + D\,\mathrm{e}^{\kappa x} \end{array}$$

On doit avoir $A = D = 0$ pour que $\varphi(x)$ soit normalisable. Compte tenu de la relation (8.93), si l'on veut avoir $(B, C) \neq 0$, il faut que S ait un pôle[9] à $k = \mathrm{i}\kappa$. Cette propriété est générale et on peut la vérifier explicitement sur le puits carré fini de la figure 8.8.

8.4 Potentiel périodique

8.4.1 Théorème de Bloch

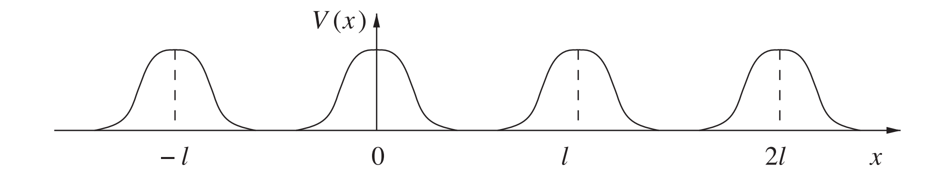


FIG. 8.11 – Potentiel périodique de période l à une dimension.

Comme dernier exemple de l'équation de Schrödinger à une dimension, considérons le cas d'un potentiel périodique de période spatiale l (figure 8.11)

$$V(x) = V(x + l) \tag{8.95}$$

Les résultats que nous obtiendrons sont d'une importance capitale en physique du solide : en effet, un électron dans un réseau cristallin est soumis à un potentiel périodique – à trois dimensions, mais les résultats obtenus ci-dessous à une dimension se généralisent à trois dimensions –, provenant des interactions de cet électron avec les ions du réseau cristallin. La périodicité du potentiel conduit à l'existence de bandes d'énergie qui, jointes au principe de Pauli, sont à la base de notre compréhension de la conductibilité électrique. Si le potentiel a la forme (8.95), le problème est invariant par toute translation $x \to x + l$, et d'après le théorème de Wigner, il existe un opérateur unitaire T_l agissant dans l'espace de Hilbert des états, en l'occurence l'espace des fonctions d'onde $L_x^{(2)}(\mathbb{R})$ tel que

$$(T_l\varphi)(x) = \varphi(x - l) \qquad T_l^\dagger = T_l^{-1} \tag{8.96}$$

9. Ou plus généralement une singularité, mais on peut montrer qu'états liés et résonances correspondent à des pôles.

Rappelons que la fonction obtenue à partir de $\varphi(x)$ par une translation de l est $\varphi(x-l)$. Comme l'opérateur T_l est unitaire, ses valeurs propres t_l sont de module un et on peut les écrire en fonction d'un paramètre q

$$t_l(q) = \mathrm{e}^{-\mathrm{i}ql} \tag{8.97}$$

Le paramètre q n'est défini qu'à un multiple entier près de $2\pi/l$: en effet, si

$$q \to q' = q + \frac{2\pi p}{l} \qquad p = 0,\, \pm 1,\, \pm 2,\, \ldots \tag{8.98}$$

la valeur de t_l est inchangée. Comme T_l commute avec le hamiltonien en raison de la périodicité (8.95) du potentiel, il est possible de diagonaliser simultanément T_l et H. Soit $\varphi_q(x)$ les fonctions propres communes de T_l et de H

$$\begin{aligned} T_l\varphi_q(x) &= t_l(q)\varphi_q(x) = \mathrm{e}^{-\mathrm{i}ql}\,\varphi_q(x) \\ H\varphi_q(x) &= E_q\varphi_q(x) \end{aligned} \tag{8.99}$$

La première de ces équations montre que

$$\varphi_q(x-l) = \mathrm{e}^{-\mathrm{i}ql}\,\varphi_q(x)$$

et nous en déduisons le *théorème de Bloch*[10] : les état stationnaires dans un potentiel périodique (8.95) sont de la forme

$$\varphi_q(x) = \mathrm{e}^{\mathrm{i}qx}\,u_{sq}(x) \qquad u_{sq}(x) = u_{sq}(x+l) \tag{8.100}$$

où $u_{sq}(x)$ est une fonction périodique de période l. L'indice s est en général nécessaire, car à une valeur de q correspondent plusieurs solutions possibles : nous verrons que s indice les bandes d'énergie. Il est facile d'écrire l'équation différentielle que vérifie $u_{sq}(x)$: comme $P = -\mathrm{i}\hbar\mathrm{d}/\mathrm{d}x$

$$\begin{aligned} P\mathrm{e}^{\mathrm{i}qx} &= \hbar q\,\mathrm{e}^{\mathrm{i}qx} \\ P\varphi_q(x) &= \mathrm{e}^{\mathrm{i}qx}\,(P+\hbar q)u_{sq}(x) \\ P^2\varphi_q(x) &= \mathrm{e}^{\mathrm{i}qx}\,(P+\hbar q)^2u_{sq}(x) \end{aligned}$$

d'où

$$H\varphi_q(x) = \mathrm{e}^{\mathrm{i}qx}\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \mathrm{i}\frac{\hbar^2 q}{m}\frac{\mathrm{d}}{\mathrm{d}x} + \frac{\hbar^2q^2}{2m} + V(x)\right]u_{sq}(x) = E_{sq}\mathrm{e}^{\mathrm{i}qx}\,u_{sq}(x)$$

soit, en divisant par $\exp(\mathrm{i}qx)$

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \mathrm{i}\frac{\hbar^2 q}{m}\frac{\mathrm{d}}{\mathrm{d}x} + \frac{\hbar^2q^2}{2m} + V(x)\right]u_{sq}(x) = E_{sq}u_{sq}(x) \tag{8.101}$$

10. Ce théorème est aussi connu sous le nom de théorème de Floquet quand on traite d'une périodicité temporelle.

La fonction d'onde dans un potentiel périodique s'obtient en résolvant (8.101) par exemple dans l'intervalle $[0, l]$ avec la condition aux limites $u_{sq}(0) = u_{sq}(l)$. La quantité $\hbar q$ a les dimensions d'une impulsion et présente effectivement certaines analogies avec une impulsion. Cependant, il ne s'agit pas d'une impulsion véritable, ne serait-ce que parce que suivant (8.98), q n'est pas unique : on appelle $\hbar q$ une *quasi-impulsion*. On remarque enfin que si le potentiel est pair, $V(x) = V(-x)$, l'équation (8.101) est inchangée dans les transformations simultanées $x \to -x,\ q \to -q$; $u_{s,-q}(x)$ est donc solution de (8.101) avec la même valeur de l'énergie, $E_{sq} = E_{s,-q}$, et tous les niveaux sont doublement dégénérés.

8.4.2 Bandes d'énergie

Examinons maintenant les propriétés des solutions de l'équation de Schrödinger (8.101) dans le potentiel périodique de la figure 8.11 : $V(x)$ est une suite de créneaux, et $V(x)$ est non nul dans des intervalles centrés autour de $x = pl$, $p = \ldots, -2, -1, 0, 1, 2 \ldots$ et s'annule dans les intervalles[11]

$$\left(p - \frac{1}{2}\right) l - \Delta x \leq x \leq \left(p - \frac{1}{2}\right) l + \Delta x \tag{8.102}$$

Dans les intervalles (8.102) où $V(x)$ s'annule, une solution $\varphi(x)$ de l'équation de Schrödinger est une superposition d'ondes planes de vecteur d'onde $\pm k$, $k = (2mE/\hbar^2)^{1/2}$. À gauche du créneau n et dans l'intervalle (8.102) pour $p = n$, $\varphi(x)$ s'écrit

$$\varphi(x) = A_n \mathrm{e}^{\mathrm{i}kx} + B_n \mathrm{e}^{-\mathrm{i}kx}$$

et à droite de ce créneau, dans l'intervalle (8.102) avec $p = n + 1$

$$\varphi(x) = A_{n+1} \mathrm{e}^{\mathrm{i}kx} + B_{n+1} \mathrm{e}^{-\mathrm{i}kx}$$

Les coefficients (A_n, B_n) sont reliés aux coefficients (A_{n+1}, B_{n+1}) suivant (8.90) par la matrice de passage M (8.92) correspondant au créneau $V(x)$

$$\begin{pmatrix} A_n \\ B_n \end{pmatrix} = \begin{pmatrix} \gamma & \delta \\ \delta^* & \gamma^* \end{pmatrix} \begin{pmatrix} A_{n+1} \\ B_{n+1} \end{pmatrix} \tag{8.103}$$

D'autre part, compte tenu du théorème de Bloch (8.100)

$$\varphi(x + l) = \mathrm{e}^{\mathrm{i}ql} \varphi(x)$$

et nous avons

$$A_{n+1}\, \mathrm{e}^{\mathrm{i}kl}\, \mathrm{e}^{\mathrm{i}kx} + B_{n+1}\, \mathrm{e}^{-\mathrm{i}kl}\, \mathrm{e}^{-\mathrm{i}kx} = \mathrm{e}^{\mathrm{i}ql} \left(A_n \mathrm{e}^{\mathrm{i}kx} + B_n \mathrm{e}^{-\mathrm{i}kx}\right)$$

11. En fait, il n'est pas nécessaire de supposer cette annulation pour obtenir les résultats qui vont suivre, mais cette supposition simplifie l'argument.

soit

$$e^{iql}\begin{pmatrix} A_n \\ B_n \end{pmatrix} = \begin{pmatrix} e^{ikl}\, A_{n+1} \\ e^{-ikl}\, B_{n+1} \end{pmatrix} = D\begin{pmatrix} A_{n+1} \\ B_{n+1} \end{pmatrix} = DM^{-1}\begin{pmatrix} A_n \\ B_n \end{pmatrix} \tag{8.104}$$

D est la matrice diagonale d'éléments $D_{11} = \exp(\,ikl)$, $D_{22} = \exp(\,-ikl)$ et

$$DM^{-1} = \begin{pmatrix} \gamma^*\, e^{ikl} & -\delta\, e^{ikl} \\ -\delta^*\, e^{-ikl} & \gamma\, e^{-ikl} \end{pmatrix} \tag{8.105}$$

L'équation (8.104) implique que (A_n, B_n) est vecteur propre de la matrice $\tilde{M} = DM^{-1}$ avec la valeur propre $\exp(iql)$, qui est de module un. L'équation aux valeurs propres λ pour la matrice $\tilde{M}$ est ($\det \tilde{M} = 1$)

$$\lambda^2 - 2\lambda \mathrm{Re}\,(\gamma^* e^{ikl}) + 1 = 0$$

et posant $x = \mathrm{Re}\,(\gamma^* \exp(\,ikl))$, les valeurs propres $\lambda_\pm$ sont données par

$$\begin{aligned} \lambda_\pm &= x \pm \sqrt{x^2 - 1} \qquad\qquad |x| > 1 \\ \lambda_\pm &= x \pm i\sqrt{1 - x^2} \qquad\qquad |x| \le 1 \end{aligned}$$

Le cas $|x| > 1$ est exclu car les racines ne peuvent pas être de module un : leur produit est égal à un et elles sont réelles. En revanche, les deux racines complexes sont bien de module un pour $|x| \le 1$; elles sont non dégénérées si $|x| < 1$ et dégénérées si $|x| = 1$.

Afin de simplifier au maximum les calculs, nous allons examiner un cas limite, le créneau en fonction delta. Les résultats que nous obtiendrons se généralisent qualitativement à tout potentiel périodique. Le potentiel périodique (8.95) est donc

$$V(x) = \sum_{p=-\infty}^{\infty} \frac{\hbar^2 g}{2m}\, \delta(x - lp) \tag{8.106}$$

On montre dans l'exercice 8.6.7 que le coefficient γ de M vaut

$$\gamma = 1 + i\,\frac{g}{2k} \tag{8.107}$$

On en déduit

$$x = \mathrm{Re}\,(\gamma^* e^{ikl}) = \cos kl + \frac{g}{2k}\,\sin kl$$

et l'équation aux valeurs propres s'écrit

$$x = \cos ql = \cos kl + \frac{g}{2k}\,\sin kl \tag{8.108}$$

Notons que q n'est pas fixé de façon unique par (8.108) : $q' = q + 2\pi p/l$ avec p entier vérifie aussi (8.108). Cette équation montre que certains intervalles en k, et donc certaines zones d'énergie puisque $E = \hbar^2 k^2/(2m)$, sont exclus parce

que le membre de droite de (8.108) peut être plus grand que un en module : ces zones sont appelées *bandes interdites*. Montrons-le explicitement dans la région $k \simeq 0$. Posons $y = kl$ et

$$f(y) = \cos y + \frac{gl}{2y} \sin y$$

Comme $f(0) = 1 + gl/2$, on voit qu'un intervalle $0 \leq y < y_0$ ou $0 \leq k < k_0$ est interdit. Supposant $gl \ll 1$ pour une estimation analytique, on trouve

$$y_0 \simeq \sqrt{gl} \quad \text{ou} \quad k_0 \simeq \sqrt{g/l}$$

Il existe d'autres zones interdites : en effet si

$$y = n\pi + \varepsilon \quad |\varepsilon| \ll 1$$

alors

$$|f(y)| \simeq 1 + \frac{gl}{2y}\,\varepsilon$$

et on met en évidence une région interdite où $|f(y)| > 1$ pour $0 < \varepsilon \ll 1$. Ces remarques permettent de tracer qualitativement la courbe $f(y)$ sur la figure 8.12. On porte conventionnellement E en fonction de q (rappelons que

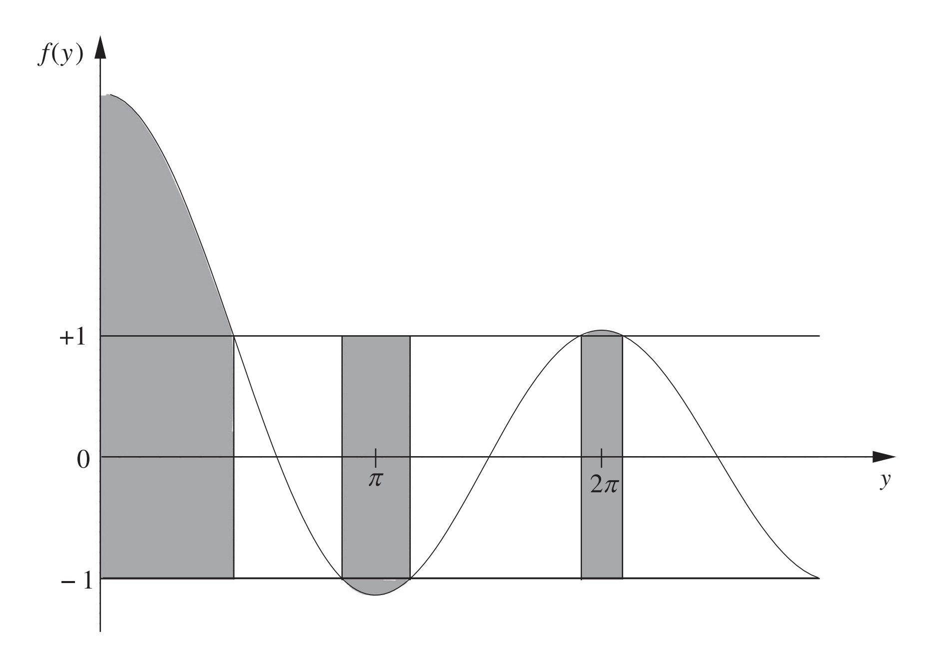


FIG. 8.12 – Solutions de (8.108).

$\hbar q$ est la quasi-impulsion), ce qui donne la figure 8.13 où l'on distingue les bandes permises indicées par s. Compte tenu de (8.98), on peut restreindre q à l'intervalle $[0, 2\pi/l]$, ou de façon équivalente à l'intervalle $[-\pi/l, \pi/l]$, qui est appelé *première zone de Brillouin.* Dans certaines régions, on peut exprimer simplement E en fonction de q. Examinons par exemple la région $k \simeq k_0$. Comme $\cos ql = 1$ pour $k = k_0$, (8.108) devient, compte tenu de $f(k_0 l) = 1$

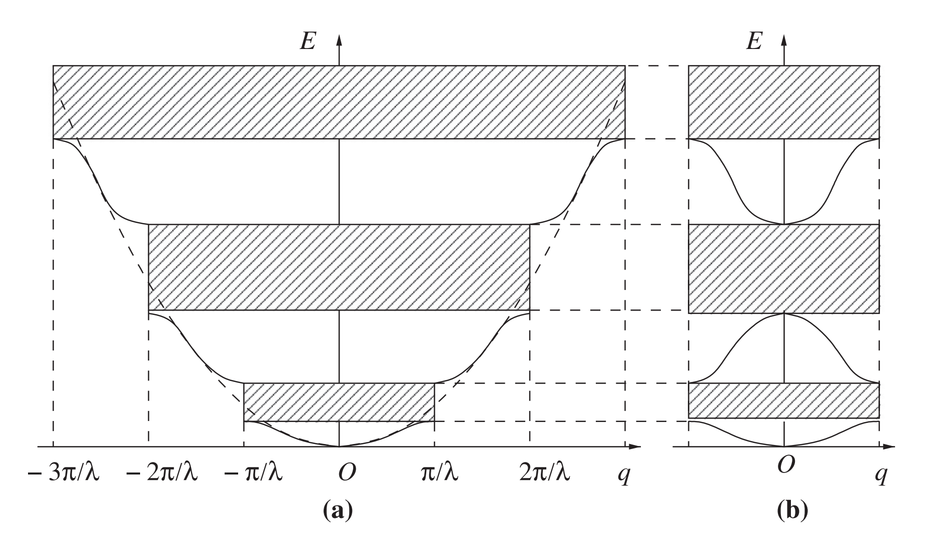


FIG. 8.13 – Bandes d'énergie : (a) q varie sans restrictions ; (b) q est limité à la première zone de Brillouin. Les zones hachurées sont les bandes interdites.

$$-\frac{1}{2}\, q^2 l^2 \simeq (k - k_0) l f'(k_0 l)$$

Ceci permet d'estimer $(E - E_0)$

$$E - E_0 = \frac{\hbar^2}{2m}\,(k^2 - k_0^2) \simeq \frac{\hbar^2 k_0 (k - k_0)}{m}$$

soit

$$E - E_0 = \frac{\hbar^2 l k_0}{2m|f'(k_0 l)|}\, q^2 = \frac{\hbar^2}{2m^*}\, q^2 \tag{8.109}$$

Au voisinage de $k = k_0$, le comportement de l'énergie est celui d'une particule de masse effective m^*

$$m^* = \frac{m|f'(k_0 l)|}{l k_0} \tag{8.110}$$

Cette masse effective joue un grand rôle dans la théorie de la conductibilité électrique : en première approximation, l'effet du réseau cristallin se traduit par un simple changement de la masse.

8.5 Mécanique ondulatoire en dimension $d = 3$

8.5.1 Généralités

Soient $\vec{R}$ et $\vec{P}$ les opérateurs position et impulsion dans l'espace à trois dimensions, de composantes[12] X_j et P_j, $j = x, y, z$. Rappelons les relations de commutation canoniques (7.45)

$$[X_j, P_k] = \mathrm{i}\hbar\, \delta_{jk}\, I \tag{8.111}$$

Les composantes de $\vec{R}$ et de $\vec{P}$ commutent si $j \neq k$. On peut donc construire l'espace des états comme produit tensoriel des espaces $L_x^{(2)}(\mathbb{R})$, $L_y^{(2)}(\mathbb{R})$ et $L_z^{(2)}(\mathbb{R})$

$$L_{\vec{r}}^{(2)}(\mathbb{R}^3) = L_x^{(2)}(\mathbb{R}) \otimes L_y^{(2)}(\mathbb{R}) \otimes L_z^{(2)}(\mathbb{R}) \tag{8.112}$$

Dans cet espace, la composante X de $\vec{R}$ sera l'opérateur

$$X \otimes I_y \otimes I_z$$

Si $\varphi_n(x)$ est une base orthonormée de $L_x^{(2)}(\mathbb{R})$, on construit une base $\varphi_{nlm}(x, y, z)$ de $L_{\vec{r}}^{(2)}(\mathbb{R}^3)$ en prenant les produits[13]

$$\varphi_{nlm}(x, y, z) = \varphi_n(x)\varphi_m(y)\varphi_l(z) \tag{8.113}$$

La construction de l'espace des états et des bases orthonormées est strictement parallèle à celle de l'espace des états de deux systèmes à deux niveaux. Nous avons observé au § 2.4.2 que le vecteur d'état le plus général de l'espace des états de systèmes à deux niveaux n'est pas en général le produit tensoriel $|\varphi_1 \otimes \varphi_2\rangle$ de deux vecteurs d'état systèmes individuels. De la même façon, une fonction $\psi(x, y, z)$ de $L_{\vec{r}}^{(2)}(\mathbb{R}^3)$ n'est pas en général un produit $\varphi(x)\chi(y)\eta(z)$, mais $\psi(x, y, z)$ peut se décomposer sur la base (8.113)

$$\psi(x, y, z) = \sum_{n,m,l} c_{nml}\, \varphi_n(x)\varphi_m(y)\varphi_l(z) \tag{8.114}$$

$$c_{nlm} = \int \mathrm{d}^3 r\, \varphi_n^*(x)\varphi_m^*(y)\varphi_l^*(z)\psi(x, y, z) \tag{8.115}$$

Il est immédiat d'écrire la généralisation à trois dimensions des formules de la section 8.1. Nous nous contentons de donner quelques exemples, en laissant au lecteur le soin d'établir les autres formules.

---

12. Les composantes de $\vec{R}$ seront notées (X, Y, Z) et celles de $\vec{r}$ (x, y, z).

13. Pour la simplicité de l'écriture, nous avons pris les mêmes fonctions de base dans les espaces (x, y, z), mais nous aurions bien sûr pu choisir trois bases différentes.

- États propres $|\vec{r}\rangle$ de $\vec{R}$

$$\vec{R}\,|\vec{r}\rangle = \vec{r}\,|\vec{r}\rangle \tag{8.116}$$

- Relation de fermeture (*cf.* (8.9))

$$\int \mathrm{d}^3 r\, |\vec{r}\rangle\langle\vec{r}| = I \tag{8.117}$$

- Amplitude de probabilité $\varphi(\vec{r})$ pour trouver une particule dans l'état $|\varphi\rangle$ au point $\vec{r}$, ou fonction d'onde de la particule

$$\varphi(\vec{r}) = \langle\vec{r}|\varphi\rangle \tag{8.118}$$

- Densité de probabilité de présence : $|\varphi(\vec{r})|^2\mathrm{d}^3r$ est la probabilité de trouver la particule dans le volume d^3r autour du point $\vec{r}$.
- Action des opérateurs $\vec{R}$ et $\vec{P}$ sur $\varphi(\vec{r})$ (*cf.* (8.14) et (8.16))

$$\boxed{\left(\vec{R}\varphi\right)(\vec{r}) = \vec{r}\varphi(\vec{r}) \qquad \left(\vec{P}\varphi\right)(\vec{r}) = -\mathrm{i}\hbar\vec{\nabla}\varphi(\vec{r})} \tag{8.119}$$

- Transformée de Fourier

$$\boxed{\tilde{\varphi}(\vec{p}) = \frac{1}{(2\pi\hbar)^{3/2}}\int \mathrm{d}^3 r\, \varphi(\vec{r})\, \mathrm{e}^{-\mathrm{i}\vec{p}\cdot\vec{r}/\hbar}} \tag{8.120}$$

On notera le facteur $(2\pi\hbar)^{-1/2}$ pour chaque dimension d'espace.

Nous avons établi au § 7.4.2 la forme générale du hamiltonien en dimension $d = 3$. Dans la suite de cette section, *nous supposerons que $\vec{A}$ est un gradient* : $\vec{A} = \vec{\nabla}\Lambda(\vec{r})$. Physiquement, cela veut dire qu'il n'y a pas de champ magnétique ; le cas d'un champ magnétique non nul sera étudié dans la section 10.3. Le hamiltonien (7.74) est simplement

$$H = \frac{\vec{P}^2}{2m} + V(\vec{R}) \tag{8.121}$$

L'équation de Schrödinger indépendante du temps[14] qui généralise (8.57) à trois dimensions est

$$\boxed{\left(-\frac{\hbar^2}{2m}\,\vec{\nabla}^2 + V(\vec{r})\right)\varphi(\vec{r}) = E\varphi(\vec{r})} \tag{8.122}$$

La généralisation du courant (8.63) est

$$\boxed{\vec{\jmath}(\vec{r},t) = \mathrm{Re}\left[\frac{\hbar}{\mathrm{i}m}\,\varphi^*(\vec{r},t)\vec{\nabla}\varphi(\vec{r},t)\right]} \tag{8.123}$$

qui vérifie l'équation de continuité (exercice 8.6.9)

$$\frac{\partial|\varphi(\vec{r},t)|^2}{\partial t} + \vec{\nabla}\cdot\vec{\jmath}(\vec{r},t) = 0 \tag{8.124}$$

14. Nous laissons au lecteur le soin d'écrire l'équation de Schrödinger dépendant du temps qui généralise (8.55) à trois dimensions.

8.5.2 Espace de phase et densité de niveaux

Dans de nombreux problèmes, il est nécessaire de savoir compter le nombre de niveaux d'énergie dans une certaine région de l'espace $(\vec{r}, \vec{p})$: cet espace est appelé *espace de phase*. Revenons au puits infini du § 8.3.3 en appelant L_x la largeur du puits. Les niveaux d'énergie sont étiquetés par un entier positif n, et nous allons nous intéresser au cas où $n \gg 1$ et où L_x est grand : les niveaux d'énergie (8.76) sont alors très resserrés et on pourra remplacer les sommes sur n par des intégrales. Soit un vecteur d'onde (8.75) $k_n = \pi(n+1)/L_x$. Calculons le nombre de niveaux d'énergie dans l'intervalle en k : $[k_n, k_n + \Delta k]$. D'après (8.77) avec $a \to L_x$, le nombre de niveaux Δn $(1 \ll \Delta n \ll n)$ dans l'intervalle $[k, k + \Delta k]$ est

$$\Delta n = \frac{L_x}{\pi}\,\Delta k \tag{8.125}$$

Au lieu des conditions d'annulation de la fonction d'onde aux points $x = 0$ et $x = L_x$, il est souvent plus commode de choisir des *conditions aux limites périodiques* : $\varphi(0) = \varphi(L_x)$, soit pour les fonctions d'onde[15]

$$\varphi_n(x) = \frac{1}{\sqrt{L_x}}\,\mathrm{e}^{\mathrm{i}k_n x} \qquad k_n = \frac{2\pi n}{L_x} \qquad n = \ldots, -2, -1, 0, 1, 2, \ldots \tag{8.126}$$

et par conséquent

$$\Delta n = \frac{L_x}{2\pi}\,\Delta k \tag{8.127}$$

À première vue, (8.127) diffère[16] de (8.125) par un facteur 1/2. Cependant, nous avons déjà observé qu'avec les fonctions d'onde (8.77) les valeurs k_n et $-k_n$ correspondent au même état physique car la substitution $k_n \to -k_n$ conduit à un simple changement de signe de la fonction d'onde. En revanche, la substitution $k_n \to -k_n$ dans (8.126) conduit à un état physique différent : la division par deux dans (8.127) est compensée par le doublement du nombre de valeurs possibles de k_n. Pour compter les niveaux d'énergie, il est équivalent de choisir les conditions aux limites périodiques ou les conditions d'annulation (voir aussi la note 17).

Passons au puits carré infini en dimension $d = 3$: les fonctions d'onde s'annulent en dehors des intervalles où $V(\vec{r}) = 0$

$$0 \le x \le L_x \qquad 0 \le y \le L_y \qquad 0 \le z \le L_z \tag{8.128}$$

15. Ce choix de fonction d'onde est parfois appelé “quantification dans une boîte”. Il évite de travailler avec les ondes planes du spectre continu, puisque les “ondes planes” (8.126) sont normalisables. Cependant, les intégrales de Fourier sont alors remplacées par des sommes de Fourier, ce qui alourdit les calculs.

16. Comme $n \gg 1$, on refait pas de différence entre n et $(n + 1)$.

Ces fonctions d'onde sont de la forme

$$\varphi_{[n_x,n_y,n_z]}(x,y,z) = \sqrt{\frac{8}{L_xL_yL_z}}\,\sin\left(\frac{\pi(n_x+1)x}{L_x}\right)\sin\left(\frac{\pi(n_y+1)y}{L_y}\right) \times \sin\left(\frac{\pi(n_z+1)z}{L_z}\right) \qquad (8.129)$$

avec $(n_x, n_y, n_z) = 0, 1, 2, \ldots$ Les valeurs correspondantes de l'énergie sont

$$E(n_x,n_y,n_z) = \frac{\hbar^2\pi^2}{2m}\left(\frac{(n_x+1)^2}{L_x^2} + \frac{(n_y+1)^2}{L_y^2} + \frac{(n_z+1)^2}{L_z^2}\right) \qquad (8.130)$$

Lorsque $L_x = L_y = L_z = L$, ces valeurs propres sont en général dégénérées (exercice 8.6.8).

Effectuons le décompte des niveaux à trois dimensions. Il sera commode d'utiliser des conditions aux limites périodiques

$$\varphi(x,y,z) = \varphi(x+L_x, y+L_y, z+L_z) \qquad (8.131)$$

Soit $\Delta\mathcal{K}$ l'élément de volume $\Delta k_x\Delta k_y\Delta k_z$ de l'espace des $\vec{k}$, tel que l'extrémité du vecteur d'onde $\vec{k}$ se trouve dans $\Delta\mathcal{K}$: les composantes (x,y,z) de ce vecteur sont comprises dans les intervalles

$$[k_x, k_x+\Delta k_x],\ [k_y, k_y+\Delta k_y],\ [k_z, k_z+\Delta k_z]$$

Le nombre de niveaux d'énergie dans $\Delta\mathcal{K}$ s'obtient en généralisant (8.127)

$$\Delta n = \left(\frac{L_x}{2\pi}\right)\Delta k_x\left(\frac{L_y}{2\pi}\right)\Delta k_y\left(\frac{L_z}{2\pi}\right)\Delta k_z = \frac{L_xL_yL_z}{(2\pi)^3}\Delta\mathcal{K} \qquad (8.132)$$

Prenant $\Delta\mathcal{K}$ infinitésimal, $\Delta\mathcal{K} = \mathrm{d}^3k$, on définit la *densité de niveaux (ou densité d'états)* $\mathcal{D}(\vec{k})$ dans l'espace des $\vec{k}$ comme suit : $\mathcal{D}(\vec{k})\mathrm{d}^3k$ est le nombre de niveaux dans le volume d^3k centré en $\vec{k}$. D'après (8.132),

$$\boxed{\mathcal{D}(\vec{k})\mathrm{d}^3k = \frac{\mathcal{V}}{(2\pi)^3}\,\mathrm{d}^3k} \qquad (8.133)$$

où $\mathcal{V} = L_xL_yL_z$ est le volume de la boîte de côtés (L_x, L_y, L_z)[17]. Compte tenu de $\vec{p} = \hbar\vec{k}$, la densité de niveaux[18] dans l'espace des $\vec{p}$ est

$$\boxed{\mathcal{D}(\vec{p}) = \frac{\mathcal{V}}{(2\pi\hbar)^3} = \frac{\mathcal{V}}{h^3}} \qquad (8.134)$$

17. Le résultat est aussi valable pour une boîte qui n'est pas un parallélépipède. Les termes correctifs sont en puissances de $(kL)^{-1}$, où L est une dimension caractéristique de la boîte. Le premier terme correctif représente un terme de surface. La différence entre conditions aux limites périodiques et conditions d'annulation, qui est un effet de surface, se traduit aussi par ce type de correction. Ces corrections sont négligeables dans une boîte suffisamment grande.

18. En toute rigueur, nous devrions utiliser une notation différente pour les différentes densités de niveaux ; nous avons utilisé la même lettre $\mathcal{D}$ dans tous les cas afin de ne pas multiplier à l'excès les notations.

Ce résultat est fondamental. On en déduit la densité de niveaux par unité d'énergie[19] ; comme $\mathcal{D}(\vec{p})$ dépend seulement de $p = |\vec{p}|$

$$\mathcal{D}(p) = \frac{4\pi\,\mathcal{V}}{(2\pi\hbar)^3}\,p^2 = \frac{\mathcal{V}}{2\pi^2\hbar^3}\,p^2 \tag{8.135}$$

La densité de niveaux par unité d'énergie $\mathcal{D}(E)$ est

$$\mathcal{D}(E) = \frac{\mathcal{V}}{2\pi^2\hbar^3}\,p^2\,\frac{\mathrm{d}p}{\mathrm{d}E} = \frac{\mathcal{V}}{2\pi^2\hbar^3}\,mp$$

ou

$$\boxed{\mathcal{D}(E) = \frac{\mathcal{V}m}{2\pi^2\hbar^3}(2mE)^{1/2}} \tag{8.136}$$

Le nombre de niveaux dans $[E, E+\mathrm{d}E]$ est $\mathcal{D}(E)\mathrm{d}E$. On peut également calculer $\mathcal{D}(E)$ à partir de $\Phi(E)$, qui est le nombre de niveaux d'énergie inférieure à E : $\mathcal{D}(E) = \Phi'(E)$ (exercice 8.6.10). La quantité $\mathcal{D}/\mathcal{V}$ est la densité de niveaux par unité de volume, qui est indépendante du volume.

En remarquant que $\mathcal{V} = \int_{\mathcal{V}} \mathrm{d}^3r$, on déduit de (8.134) que le nombre de niveaux dans $\mathrm{d}^3r\,\mathrm{d}^3p$ est

$$\boxed{\mathrm{d}N = \frac{\mathrm{d}^3r\,\mathrm{d}^3p}{(2\pi\hbar)^3} = \frac{\mathrm{d}^3r\,\mathrm{d}^3p}{h^3}} \tag{8.137}$$

$\mathrm{d}^3r\,\mathrm{d}^3p$ est un volume infinitésimal de l'espace de phase $(\vec{r}, \vec{p})$. On peut interpréter (8.137) de la façon suivante : *le volume élémentaire de l'espace de phase est h^3, dans le sens où l'on peut placer un niveau d'énergie par volume élémentaire.* L'inégalité de Heisenberg permet de le comprendre : si une particule est confinée dans un intervalle Δx, son impulsion $p \sim h/\Delta x$. Ce résultat s'exprime sous forme imagée : alors qu'une particule classique dont l'état est défini par sa position $\vec{r}$ et son impulsion $\vec{p}$ occupe un point $(\vec{r}, \vec{p})$ dans l'espace de phase, une particule quantique doit occuper au minimum un volume $\sim h^3$.

Les résultats (8.134) ou (8.136) sont d'une importance capitale en mécanique statistique quantique : la probabilité qu'un système à l'équilibre thermique ait une énergie E (voir (1.12) et la note 16 du chapitre 1) est

$$\mathsf{p}(E) = \mathcal{N}\,\mathcal{D}(E)\,\mathrm{e}^{-\beta E}$$

où $\mathcal{N}$ est une constante de normalisation fixée par

$$\int \mathrm{d}E\,\mathsf{p}(E) = 1$$

19. Avec des conditions d'annulation de la fonction d'onde, il s'introduirait un facteur 1/8 dans (8.133) pour tenir compte du fait que les composantes de $\vec{k}$ sont positives. Le résultat final serait toutefois identique, à cause du facteur 1/2 de différence entre (8.125) et (8.127) : $(1/2)^3 = 1/8$.

8.5.3 Règle d'or de Fermi

La notion de densité de niveaux va nous permettre de montrer une des formules les plus importantes de la physique quantique, la *règle d'or de Fermi*, qui permet de calculer les probabilités de transition vers les états de diffusion, aussi appelés états du continu, car appartenant au spectre continu d'un hamiltonien, dans le cas présent le hamiltonien H_0 (8.138). Considérons un système physique régi par le hamiltonien $H(t)$ dépendant du temps

$$H(t) = H_0 + W(t) \tag{8.138}$$

où H_0 est indépendant du temps et de spectre supposé connu : les valeurs propres de l'énergie sont E_n et les vecteurs propres $|n\rangle$

$$H_0|n\rangle = E_n|n\rangle \tag{8.139}$$

Nous nous proposons de résoudre le problème suivant : au temps $t = 0$, le système est dans l'état initial $|\psi(0)\rangle = |i\rangle$, état propre de H_0 d'énergie E_i, et nous voulons calculer la probabilité $\mathsf{p}_{i\to f}(t)$ de le trouver au temps t dans l'état propre $|f\rangle$ de H_0 d'énergie E_f. Pour ce faire, nous devons obtenir le vecteur d'état $|\psi(t)\rangle$ du système au temps t, puisque

$$\mathsf{p}_{i\to f}(t) = |\langle f|\psi(t)\rangle|^2 \quad \text{avec} \quad |\psi(t=0)\rangle = |i\rangle \tag{8.140}$$

Nous avons déjà rencontré ce problème dans un cas simple : nous avons calculé dans la section 5.3 la probabilité de transition d'un niveau à l'autre de la molécule d'ammoniac plongée dans un champ électromagnétique oscillant. Le hamiltonien (8.138) généralise (5.54), H_0 étant l'analogue de (5.55). Nous suivrons la méthode du § 5.3.3 en l'adaptant à un nombre de niveaux quelconque. Comme au § 5.3.3, il est commode de passer dans le point de vue de l'interaction, et suivant (4.44), de définir le vecteur d'état $|\tilde{\psi}(t)\rangle$ et l'interaction $\tilde{W}(t)$ dans ce point de vue

$$|\tilde{\psi}(t)\rangle = \mathrm{e}^{\,\mathrm{i}H_0t}|\psi(t)\rangle \qquad \tilde{W}(t) = \mathrm{e}^{\,\mathrm{i}H_0t}\,W(t)\,\mathrm{e}^{-\mathrm{i}H_0t} \tag{8.141}$$

Soit $\gamma_n(t)$ les composantes de $|\tilde{\psi}(t)\rangle$ dans la base $|n\rangle$

$$|\tilde{\psi}(t)\rangle = \sum_n \gamma_n(t)\,|n\rangle \tag{8.142}$$

L'équation d'évolution

$$\mathrm{i}\hbar\frac{\mathrm{d}|\tilde{\psi}(t)\rangle}{\mathrm{d}t} = \tilde{W}(t)|\tilde{\psi}(t)\rangle \tag{8.143}$$

s'écrit en terme des composantes $\gamma_n(t)$. Compte tenu des équations (8.141) à (8.143), on obtient pour les coefficients γ_n le système d'équations différentielles

$$\begin{aligned}\mathrm{i}\hbar\dot{\gamma}_n(t) &= \sum_l \langle n|\tilde{W}(t)|l\rangle\gamma_l(t) = \sum_l \langle n|W(t)|l\rangle \mathrm{e}^{\mathrm{i}(E_n-E_l)t/\hbar}\gamma_l(t)\\ &= \sum_l W_{nl}(t)\,\mathrm{e}^{\mathrm{i}\,\omega_{nl}t}\,\gamma_l(t)\end{aligned} \tag{8.144}$$

avec

$$W_{nl}(t) = \langle n|W(t)|l\rangle \qquad \omega_{nl} = \frac{E_n - E_l}{\hbar} \tag{8.145}$$

Ces équations sont exactes, mais ne sont pas solubles analytiquement, sauf cas particulier. Il faut avoir recours à des approximations : nous allons suivre une méthode appelée la *théorie des perturbations dépendant du temps*. Il est commode d'introduire un paramètre réel λ, $0 \le \lambda \le 1$, qui multiplie la perturbation W : $W \to \lambda W$, ce qui permet de faire varier artificiellement la perturbation[20]. La théorie des pertubations consiste à obtenir une solution approchée de l'équation de Schrödinger sous forme d'un développement en puissances de λ et on rétablit $\lambda = 1$ à la fin des calculs. Dans ce qui suit, nous allons nous contenter du premier ordre[21] en λ. Au temps $t = 0$, le système est supposé dans l'état $|i\rangle$

$$\gamma_n(0) = \delta_{ni}$$

Écrivons

$$\gamma_n(t) = \delta_{ni} + \gamma_n^{(1)}(t)$$

Lorsque t est suffisamment petit, $|\gamma_n^{(1)}(t)| \ll 1$ car le système n'a pas eu le temps d'évoluer de façon appréciable. L'équation (8.145) devient en introduisant le paramètre λ

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left(\delta_{ni} + \gamma_n^{(1)}(t)\right) = \sum_l \lambda W_{nl}(t)\left[\delta_{li} + \gamma_l^{(1)}(t)\right]\mathrm{e}^{\mathrm{i}\,\omega_{nl}t}$$

On observe que $\gamma_l^{(1)}(t)$ est d'ordre λ, et que le terme $\sum_l \lambda W_{nl}(t)\gamma_l^{(1)}(t)$ sera donc d'ordre λ^2. Ce terme est négligeable au premier ordre en λ, ce qui donne en rétablissant $\lambda = 1$

$$\mathrm{i}\hbar\dot{\gamma}_n^{(1)}(t) \simeq W_{ni}(t)\,\mathrm{e}^{\mathrm{i}\,\omega_{ni}\,t} \tag{8.146}$$

Un cas particulier important est celui du potentiel oscillant

$$W(t) = A\,\mathrm{e}^{-\mathrm{i}\,\omega t} + A^\dagger \mathrm{e}^{\mathrm{i}\,\omega t} \tag{8.147}$$

où A est un opérateur. C'est par exemple ce type de potentiel qui décrit l'interaction d'un atome avec un champ électromagnétique oscillant

$$\mathcal{E}(t) = \mathcal{E}_0\,\mathrm{e}^{-\mathrm{i}\,\omega t} + \mathcal{E}_0^*\,\mathrm{e}^{\mathrm{i}\,\omega t}$$

Si nous nous intéressons comme au chapitre 5 à la transition $i \to f$ vers un niveau final $|f\rangle$ bien déterminé, l'amplitude de probabilité $\langle f|\psi(t)\rangle$ est donnée à un facteur de phase près par $\gamma_f(t) \simeq \gamma_f^{(1)}(t)$ qui est solution de l'équation différentielle (8.146)

$$\mathrm{i}\hbar\dot{\gamma}_f^{(1)}(t) = A_{fi}\,\mathrm{e}^{-\mathrm{i}(\omega-\omega_0)t} + A_{if}^*\,\mathrm{e}^{\mathrm{i}(\omega+\omega_0)t} \tag{8.148}$$

20. Si la perturbation est due à l'interaction avec un champ extérieur, on peut la faire varier en ajustant ce champ.

21. La complexité des formules augmente très rapidement avec la puissance de λ.

avec $\omega_0 = \omega_{fi} = (E_f - E_i)/\hbar$. Cette équation différentielle s'intègre immédiatement car les coefficients $A_{fi} = \langle f|A|i\rangle$ sont indépendants du temps

$$\gamma_f^{(1)}(t) = \frac{1}{\hbar}\left[A_{fi}\frac{\mathrm{e}^{-\mathrm{i}(\omega-\omega_0)t}-1}{\omega-\omega_0} - A_{if}^*\frac{\mathrm{e}^{\mathrm{i}(\omega+\omega_0)t}-1}{\omega+\omega_0}\right] \tag{8.149}$$

Cette amplitude de probabilité sera importante si $\omega \simeq \pm\omega_0$, c'est-à-dire, comme au chapitre 5, à la résonance. Dans le cas $\omega \simeq \omega_0$,

$$E_f \simeq E_i + \hbar\omega$$

et le système absorbe une énergie $\hbar\omega$: s'il s'agit de l'interaction avec une onde électromagnétique, il absorbe un photon d'énergie $\hbar\omega$. Dans le cas $\omega \simeq -\omega_0$,

$$E_f \simeq E_i - \hbar\omega$$

et le système fournit une énergie $\hbar\omega$, par exemple en émettant un photon[22] d'énergie $\hbar\omega$. Pour fixer les idées, examinons le premier cas. La probabilité de transition $\mathsf{p}_{i\to f}(t)$ vaut

$$\mathsf{p}_{i\to f}(t) = |\gamma_f^{(1)}(t)|^2 = \frac{1}{\hbar^2}\,|A_{fi}|^2 t^2 f(\omega-\omega_0;t) \tag{8.150}$$

où la fonction f a été définie en (5.66)

$$f(\omega-\omega_0;t) = \frac{\sin^2[(\omega-\omega_0)t/2]}{[(\omega-\omega_0)t/2]^2} \simeq \frac{2\pi}{t}\,\delta(\omega-\omega_0) \tag{8.151}$$

On retrouve les résultats du § 5.3.3 dans un cas plus général. Compte tenu des approximations, une condition nécessaire pour la validité de (8.150) est que $\mathsf{p}_{i\to f}(t) \ll 1$. Cependant, il est en général impossible d'isoler une transition vers un état f particulier, et on s'intéresse le plus souvent à une transition vers un ensemble d'états finaux d'énergie voisine, transition à laquelle correspond une probabilité par unité de temps Γ

$$\Gamma = \sum_f \Gamma_{i\to f}$$

La sommation sur f est équivalente à une intégration sur l'énergie si l'on tient compte de la densité de niveaux $\mathcal{D}(E)$

$$\sum_f \to \int \mathrm{d}E\,\mathcal{D}(E)$$

Par exemple, si l'état final correspond à celui d'une particule et si $|A_{fi}|^2$ est isotrope, la densité de niveaux sera donnée par (8.136). Si $|A_{fi}|^2$ n'est pas

22. Voir, cependant, la mise en garde de la note 27 du chapitre 1.

isotrope, et dépend par exemple de la direction de l'impulsion $\vec{p}$ de la particule finale, on utilisera

$$\mathcal{D}(E) = \frac{\mathcal{V}m}{2\pi^2\hbar^3}(2mE)^{1/2}\,\frac{\mathrm{d}\Omega}{4\pi}$$

où Ω repère la direction de $\vec{p}$. Compte tenu de (8.150) et (8.151), nous obtenons une probabilité de transition par unité de temps Γ

$$\begin{aligned}\Gamma &= \frac{1}{\hbar^2}\int \mathrm{d}E\,|A_{fi}|^2\,\mathcal{D}(E)\,t\,\frac{\sin^2[(\omega-\omega_0)t/2]}{[(\omega-\omega_0)t/2]^2} \\ &\simeq \frac{1}{\hbar}\int \mathrm{d}E\,|A_{fi}|^2\,\mathcal{D}(E)\,2\pi\,\delta(E-(E_i+\hbar\omega))\end{aligned}$$

En effectuant l'intégrale, on obtient la règle d'or de Fermi

$$\boxed{\Gamma = \frac{2\pi}{\hbar}\,|A_{fi}|^2\,\mathcal{D}(E_f) \qquad E_f = E_i + \hbar\omega} \tag{8.152}$$

Cette équation est aussi valable dans le cas d'émission d'énergie, en prenant $E_f = E_i - \hbar\omega$, et aussi pour un potentiel W constant, avec $E_f = E_i$ (exercice 8.6.11).

Les conditions de validité du calcul sont les suivantes.

- Il est nécessaire que la probabilité de trouver le système dans l'état initial (i) soit proche de un, soit

$$\sum_{f\neq i}\mathsf{p}_{i\to f}(t) \ll 1 \;\;\text{ou, en termes de } \Gamma_{i\to f}\;:\;\left(\sum_{f\neq i}\Gamma_{i\to f}\right)t \ll 1$$

  ce qui implique que t doit être suffisamment court : $t \ll \tau_2$.
- Dans l'intégrale sur l'énergie E, on doit pouvoir remplacer $f(\omega-(E-E_i)/\hbar;t)$ par une fonction delta

$$\int \mathrm{d}E\,g(E)f\left(\omega-\frac{E-E_i}{\hbar};t\right) \to \int \mathrm{d}\omega\,g(E)\,\frac{2\pi}{t}\,\delta(E-\hbar\omega_0) = \frac{2\pi}{t}\,g(E_f)$$

  Si ΔE_1 est l'intervalle caractéristique de variation de $g(E) = |A_{fi}|^2\,\mathcal{D}(E)$, $\tau_1 = \hbar/\Delta E_1$ doit être petit par rapport à t : $t \gg \tau_1$.

En résumé, il faut que t vérifie l'encadrement $\tau_1 \ll t \ll \tau_2$.

Une application importante de la règle d'or de Fermi est la désintégration d'un état instable i (état excité d'un atome ou d'un noyau, particule instable...) vers un continuum d'états f. La perturbation est alors indépendante du temps et $E_f \simeq E_i$ dans (8.152). Pour des temps suffisamment courts, la probabilité de trouver le système dans l'état initial instable i est

$$\mathsf{p}_{ii}(t) = 1 - \Gamma t \simeq \mathrm{e}^{-\Gamma t} \qquad t \ll \tau_2 \tag{8.153}$$

et il est tentant d'identifier Γ à l'inverse de la vie moyenne τ : $\Gamma = \hbar/\tau$. Le calcul que nous venons d'effectuer ne nous permet pas cette identification,

car il n'est pas *a priori* valable quel que soit t. On peut cependant justifer la loi exponentielle (8.153) pour des temps longs, grâce à une méthode due à Wigner et Weisskopf décrite dans l'appendice B. Cette méthode montre que la dispersion en énergie ΔE sur l'énergie E_f des états finaux est $\Delta E = \hbar/2\tau = \hbar\Gamma/2$.

8.6 Exercices

8.6.1 Inégalités de Heisenberg

1. Soit $\varphi(x)$, une fonction de carré sommable normalisée à l'unité et $I(\alpha)$, la quantité positive ou nulle

$$I(\alpha) = \int_{-\infty}^{\infty} \mathrm{d}x \left| x\varphi(x) + \alpha \frac{\mathrm{d}\varphi}{\mathrm{d}x} \right|^2 \geq 0$$

α étant un nombre réel. En intégrant par parties, montrer que

$$I(\alpha) = \langle X^2 \rangle - \alpha + \alpha^2 \langle K^2 \rangle$$

où $K = -\mathrm{i}\,\mathrm{d}/\mathrm{d}x$ et

$$\langle X^2 \rangle = \int_{-\infty}^{\infty} \mathrm{d}x\, x^2 |\varphi(x)|^2 \qquad \langle K^2 \rangle = -\int_{-\infty}^{\infty} \mathrm{d}x\, \varphi^*(x) \frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2}$$

En déduire

$$\langle X^2 \rangle \langle K^2 \rangle \geq \frac{1}{4}$$

2. Comment faut-il modifier le raisonnement de la question précédente pour obtenir l'inégalité de Heisenberg

$$\Delta x\, \Delta k \geq \frac{1}{2} \ ?$$

Montrer que $\Delta x\, \Delta k = 1/2$ implique que $\varphi(x)$ est une gaussienne

$$\varphi(x) \propto \exp\left(-\frac{1}{2} \sigma^2 x^2 \right)$$

.

8.6.2 Étalement du paquet d'ondes

1. Montrer que $[P^2, X] = -2\mathrm{i}\hbar\, P$

2. Soit $\langle X^2 \rangle(t)$ la position quadratique moyenne dans l'état $|\varphi(t)\rangle$

$$\langle X^2 \rangle(t) = \langle \varphi(t) | X^2 | \varphi(t) \rangle$$

Montrer que

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle X^2\rangle(t) &= \frac{1}{m}\langle PX + XP\rangle \\ &= \frac{\mathrm{i}\hbar}{m}\int_{-\infty}^{\infty}\mathrm{d}x\, x\left[\varphi\frac{\partial\varphi^*}{\partial x} - \varphi^*\frac{\partial\varphi}{\partial x}\right]\end{aligned}$$

Ces résultats sont-ils valables si le potentiel $V(x) \neq 0$?

3. Montrer que si la particule est libre $(V(x) = 0)$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle X^2\rangle(t) = \frac{2}{m^2}\langle P^2\rangle = 2v_1^2 = \mathrm{cste}$$

4. Déduire de ces résultats

$$\langle X^2\rangle(t) = \langle X^2\rangle(t=0) + \xi_0 t + v_1^2 t^2 \qquad \xi_0 = \left.\frac{\mathrm{d}\langle X^2\rangle}{\mathrm{d}t}\right|_{t=0}$$

ainsi que l'expression de $(\Delta x(t))^2$

$$(\Delta x(t))^2 = (\Delta x(t=0))^2 + [\xi_0 - 2v_0\langle X\rangle(t=0)]t + (v_1^2 - v_0^2)t^2$$

avec $v_0 = \langle P/m\rangle = \mathrm{cste}$.

8.6.3 Paquet d'ondes gaussien

1. On suppose que la fonction $A(k)$ de (8.44) est une gaussienne

$$A(k) = \frac{1}{(\pi\sigma^2)^{1/4}}\exp\left[-\frac{(k-\overline{k})^2}{2\sigma^2}\right]$$

Montrer que

$$\int |A(k)|^2\mathrm{d}k = 1, \qquad \Delta k = \frac{1}{\sqrt{2}}\,\sigma$$

et que la fonction d'onde $\varphi(x, t=0)$ vaut

$$\varphi(x, t=0) = \frac{\sigma^{1/2}}{\pi^{1/4}}\exp\left[\mathrm{i}\overline{k}x - \frac{1}{2}\sigma^2 x^2\right]$$

Tracer la courbe représentative de $|\varphi(x, t=0)|^2$. Quelle est la largeur de cette courbe ? Identifier la dispersion Δx et vérifier que $\Delta x\,\Delta k = 1/2$.

2. Calculer $\varphi(x,t)$. Montrer que si $\hbar\sigma^2 t/m \ll 1$, on a

$$\varphi(x,t) = \exp\left(\frac{\mathrm{i}\hbar\overline{k}^2}{2m}\,t\right)\varphi(x - v_g t, 0) \qquad v_g = \frac{\hbar\overline{k}}{m}$$

3. Calculer exactement $\varphi(x,t)$

$$\varphi(x,t) = \left(\frac{1}{\pi\sigma^2}\right)^{1/4}\sigma'\exp\left[\mathrm{i}\overline{k}x - \mathrm{i}\omega(\overline{k})t - \frac{1}{2}\sigma'^2(x - v_g t)^2\right]$$

avec

$$\frac{1}{\sigma'^2} = \frac{1}{\sigma^2} + \frac{\mathrm{i}\hbar t}{m}$$

et en déduire $|\varphi(x,t)|^2$. Montrer que

$$\Delta x^2(t) = \frac{1}{2\sigma^2}\left(1 + \frac{\hbar^2\sigma^4 t^2}{m^2}\right)$$

Donner l'interprétation physique du résultat.

4. Un neutron sort d'un réacteur nucléaire avec une longueur d'onde de 0.1 nm. On suppose que sa fonction d'onde à $t = 0$ est un paquet d'ondes gaussien de largeur $\Delta x = 1$ nm. Au bout de combien de temps la largeur du paquet d'ondes aura-t-elle doublé ? Quelle distance aura parcouru le neutron ?

8.6.4 Heuristique de l'inégalité de Heisenberg

1. Si l'électron émis dans la désintégration β du neutron

$$n \to p + e^- + \overline{\nu}_e$$

était initialement enfermé dans le neutron dont le rayon $\simeq 0.5\,\mathrm{fm}$, quelle serait son énergie cinétique ? Quelle conclusion peut-on en tirer ?

2. Une particule quantique de masse m se déplace sur l'axe des x dans le potentiel harmonique

$$V(x) = \frac{1}{2}m\omega^2\, x^2$$

Utiliser l'inégalité de Heisenberg pour estimer l'énergie de son état fondamental.

8.6.5 Potentiel de Lennard-Jones pour l'hélium

1. L'énergie potentielle de deux atomes séparés par une distance r est souvent bien représentée par le potentiel de Lennard-Jones

$$V(r) = \varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - 2\left(\frac{\sigma}{r}\right)^{6}\right]$$

où ε et σ sont des paramètres ayant respectivement les dimensions d'une énergie et d'une longueur. Calculer la position r_0 du minimum du potentiel et tracer qualitativement $V(r)$. Montrer qu'au voisinage de $r = r_0$

$$V(r) \simeq -\varepsilon\left[1 - 36\left(\frac{r - r_0}{r_0}\right)^2\right] = \frac{1}{2}m\omega^2(r - r_0)^2 + V_0$$

2. Dans le cas de l'hélium, $\varepsilon \simeq 10^{-3}\,\mathrm{eV}$ et $r_0 \simeq 0.3\,\mathrm{nm}$. Calculer la fréquence de vibration ω et l'énergie $\hbar\omega/2$ du niveau fondamental. Pourquoi

l'hélium reste-t-il liquide même si la température $T \to 0$? Le raisonnement vaut-il pour les deux isotopes ^{3}He et ^{4}He ?

3. Pour l'hydrogène, $\varepsilon \simeq 4\,\text{eV}$. Pourquoi l'hydrogène se solidifie-t-il à basse température ? Que pensez-vous des gaz rares : argon, néon, etc. ?

8.6.6 Marche de potentiel et retard à la réflexion

1. Cas $E > V_0$. On suppose que les particules sont envoyées de la gauche et arrivent sur la marche de potentiel : dans la région II, on trouve uniquement des particules se propageant vers la droite. Définissons

$$k' = \sqrt{\frac{2m(E - V_0)}{\hbar^2}}$$

Les fonctions d'onde sont dans les régions I et II

$$\begin{array}{lll} \text{I} & \varphi(x) & = \mathrm{e}^{\mathrm{i}kx} + B\,\mathrm{e}^{-\mathrm{i}kx} \\ \text{II} & \varphi(x) & = C\,\mathrm{e}^{\mathrm{i}k'x} \end{array}$$

Soit R le coefficient de réflexion et $T = 1 - R$ le coefficient de transmission. Montrer que

$$R = \left(\frac{k - k'}{k + k'}\right)^2 \qquad T = 1 - R = \frac{4kk'}{(k + k')^2}$$

Pourquoi n'a-t-on pas $T = |C|^2$? Vérifier que la conservation du courant est

$$\frac{\hbar k}{m} = \frac{\hbar k}{m}|B|^2 + \frac{\hbar k'}{m}|C|^2$$

2. Cas $E < V_0$. L'équation (8.74) donne le coefficient B de l'onde réfléchie quand une onde incidente $\exp(\mathrm{i}kx)$ arrive sur une marche de potentiel avec une énergie $E = \hbar^2 k^2/(2m) < V_0$, V_0 étant la hauteur de la marche de potentiel. Montrer que $|B| = 1$, $B = \exp(-\mathrm{i}\phi)$. Déterminer ϕ et $\mathrm{d}\phi/\mathrm{d}E$.

3. On suppose que l'onde incidente est un paquet d'ondes du type (8.44)

$$\varphi(x,t) = \int \frac{\mathrm{d}k}{\sqrt{2\pi}}\, A(k) \exp(\mathrm{i}kx - \mathrm{i}\omega(k)t)$$

Quel sera le paquet d'ondes réfléchi ? En déduire que la réflexion se fait avec un retard

$$\tau = -\hbar \frac{\mathrm{d}\phi}{\mathrm{d}E} > 0$$

8.6.7 Potentiel en fonction δ

On considère un potentiel à une dimension de la forme

$$V(x) = \frac{\hbar^2 g}{2m}\,\delta(x)$$

m étant la masse de la particule soumise au potentiel. Ce potentiel donne parfois une approximation commode. Par exemple, il peut représenter une barrière de potentiel de largeur a et de hauteur V_0, dans la limite où $a \to 0$ et $V_0 \to \infty$, le produit $V_0 a$ restant constant et égal à $\hbar^2 g/(2m)$. Dans le cas d'une barrière (potentiel répulsif), $g > 0$, mais on peut aussi modéliser un puits (potentiel attractif), auquel cas $g < 0$.

1. Montrer que g a pour dimension l'inverse d'une longueur.

2. La fonction $\varphi(x)$ obéit à l'équation de Schrödinger

$$\left[-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + g\,\delta(x)\right]\varphi(x) = \frac{2mE}{\hbar^2}\,\varphi(x)$$

Montrer que la dérivée de $\varphi(x)$ vérifie au voisinage de $x = 0$

$$\varphi'(0^+) - \varphi'(0^-) = g\,\varphi(0) \qquad \varphi(0^\pm) = \lim_{\varepsilon\to 0^\pm} \varphi(\varepsilon)$$

Supposant $g < 0$, montrer qu'il existe un état lié et un seul. Déterminer son énergie et la fonction d'onde correspondante. Montrer que l'on retrouve les résultats en prenant la limite du puits carré avec $V_0 a \to \hbar^2|g|/(2m)$ et $a \to 0$.

3. *Modèle pour une molécule diatomique.* Supposant toujours $g < 0$, on modélise très grossièrement le potentiel auquel est soumis un électron d'une molécule diatomique par

$$V(x) = \frac{\hbar^2 g}{2m}\left[\delta(x+l) + \delta(x-l)\right]$$

La droite des noyaux est prise comme axe des x et les deux noyaux sont situés en $x = -l$ et $x = +l$. Montrer que l'on peut classer les solutions de l'équation de Schrödinger en solutions paires et impaires. Si la fonction d'onde est paire, montrer qu'il existe un seul état lié donné par

$$\kappa = \frac{|g|}{2}\,(1 + \mathrm{e}^{-2\kappa l}) \qquad \kappa = \sqrt{\frac{2m|E|}{\hbar^2}}$$

Tracer qualitativement sa fonction d'onde. Si la fonction d'onde est impaire, déterminer l'équation donnant l'énergie des états liés

$$\kappa = \frac{|g|}{2}\,(1 - \mathrm{e}^{-2\kappa l})$$

Existe-t-il toujours un état lié ? Sinon, quelle condition faut-il imposer ? Tracer qualitativement la fonction d'onde lorsqu'il y a un état lié.

4. *Puits double et effet tunnel.* On reprend la question précédente en supposant que $\kappa l \gg 1$. Montrer que les deux états liés constituent un système à deux niveaux dont le hamiltonien est

$$H = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix}$$

et relier A à $\sqrt{T}$, où T est le coefficient de transmission par effet tunnel entre les deux puits.

5. *Barrière de potentiel.* On s'intéresse maintenant au cas $g > 0$, qui modélise une barrière de potentiel. Calculer la matrice de passage et donner l'expression du coefficient de transmission.

6. *Potentiel périodique.* Un électron se déplaçant dans un cristal à une dimension est soumis à un potentiel périodique de période l que l'on modélise par (8.106). Pour fixer les idées, on prendra $g > 0$. Utiliser les conditions sur $\varphi'(x)$ pour obtenir directement (8.108)

$$\cos ql = \cos kl + \frac{g}{2k} \sin kl$$

Montrer qu'il existe des régions interdites pour l'énergie. Tracer qualitativement l'énergie E_q en fonction de q.

8.6.8 Niveaux d'énergie du puits cubique infini en dimension $d = 3$

Déterminer l'énergie des six premiers niveaux d'énergie du puits cubique infini en fonction du côté L du cube ainsi que leur dégénérescence.

8.6.9 Courant de probabilité à trois dimensions

Montrer l'équation de continuité

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{\jmath} = 0 \qquad \rho = |\varphi(\vec{r}, t)|^2$$

pour le courant (8.123).

8.6.10 Densité de niveaux

1. Calculer la densité de niveaux d'énergie $\mathcal{D}(E)$ en dimension $d = 2$. Montrer qu'elle est indépendante de E.

2. Calculer directement le nombre de niveaux $\Phi(E)$ dont l'énergie est inférieure à E, en comptant le nombre de niveaux possibles dans une sphère de rayon $|p| = \sqrt{2mE}$ dans l'espace des impulsions, et en prenant garde aux conditions aux limites. Retrouver l'expression (8.136) de $\mathcal{D}(E)$ par

$$\mathcal{D}(E) = \frac{\mathrm{d}\Phi(E)}{\mathrm{d}E}$$

8.6.11 Règle d'or de Fermi

1. *Comparaison avec la formule de Rabi.* Dans un système à deux niveaux, la formule de Rabi (5.62) donne exactement la probabilité de transition entre

deux niveaux sous l'effet d'une perturbation harmonique, par exemple

$$\mathsf{p}_{+\to -}(t) = \frac{\omega_1^2}{\Omega^2}\sin^2\frac{\Omega t}{2} \qquad \Omega^2 = \left[(\omega-\omega_0)^2+\omega_1^2\right]^{1/2}$$

Montrer que la formule approchée (8.150) s'obtient comme la limite de la formule de Rabi si

- $|\omega-\omega_0| \gg \omega_1$, c'est-à dire loin de la résonance
- ou si $\omega_1 t \ll 1$, c'est-à-dire pour des temps assez courts.

2. *Potentiel constant.* Donner l'expression de l'amplitude (8.149) $\gamma^{(1)}(t)$ et de la probabilité de transition par unité de temps Γ lorsque le potentiel $W(t)$ de (8.147) est indépendant du temps.

8.6.12 Étude de l'expérience de Stern-Gerlach

1. *Étude classique.* Les notations sont celles du § 3.2.2. La trajectoire des atomes d'argent (figure 3.8) est supposée contenue dans le plan de symétrie yOz et voisine de l'axe Oy. Montrer que $\partial B_z/\partial x|_{x=0} = 0$ et que $\partial B_z/\partial y = 0$ si l'on néglige les effets de bord. Montrer qu'une forme approchée du champ magnétique vérifiant les équations de Maxwell dans l'entrefer de l'aimant au voisinage de $x = 0$ et $z = 0$ est

$$\vec{B} = B_0\hat{z} + b(z\hat{z} - x\hat{x})$$

où $b = \partial B_z/\partial z|_{z=0}$. L'expression classique de la force est $\vec{F} = -\vec{\nabla}(\vec{\mu}\cdot\vec{B})$. En déduire les composantes F_x, F_y et F_z. Montrer que sous l'effet de B_0, le moment magnétique $\vec{\mu}$ précesse autour de l'axe Oz avec une fréquence $\omega = |\gamma B_0|$, où γ est le facteur gyromagnétique, et que si $1/\omega$ est très petit par rapport au temps de transit de l'atome dans l'entrefer de l'aimant, alors la composante μ_x donne une force moyenne nulle : tout se passe comme si le moment magnétique était soumis à une force effective $\vec{F} = b\mu_z\hat{z}$.

2. *Données numériques.* Les atomes d'argent de masse $m = 1.8\times 10^{-27}$ kg sortent du four avec une vitesse $v \simeq 500\,\mathrm{m.s^{-1}}$ et une dispersion des vitesses $\Delta v \sim 10\,\mathrm{m.s^{-1}}$. Les fentes collimatrices ont une hauteur $\Delta z = 10^{-4}$ m, la longueur de l'entrefer est $L = 5\times 10^{-2}$ m, le champ magnétique $B_0 = 1$ T et $b = 10^4\,\mathrm{T.m^{-1}}$. Montrer que l'écart δ entre les deux trajectoires correspondant à $S_z = \hbar/2$ et $S_z = -\hbar/2$ à la sortie de l'entrefer vaut

$$\delta = \frac{\mu b}{m}\left(\frac{L}{v}\right)^2$$

Évaluer numériquement δ. Calculer le produit $\Delta z\Delta p_z$ et montrer que $\Delta z\Delta p_z \gg \hbar$. On peut donc traiter les trajectoires des atomes de façon classique.

3. *Description quantique.* Soit $\varphi_\pm(\vec{r}, t)$, la fonction d'onde d'un atome dont le spin est dans l'état $|\pm\rangle$. Montrer que $\varphi_\pm$ vérifie l'équation de Schrödinger

$$\mathrm{i}\hbar \frac{\partial \varphi_\pm}{\partial t} = \left(-\frac{\hbar^2}{2m} \nabla^2 \mp \mu B\right) \varphi_\pm$$

On définit la position moyenne $\langle \vec{r}_\pm \rangle(t)$ et l'impulsion moyenne $\langle \vec{p}_\pm \rangle(t)$ des paquets d'ondes $\varphi_\pm(\vec{r}, t)$ par

$$\begin{aligned} \langle \vec{r}_\pm \rangle(t) &= \int \mathrm{d}^3 r \, \vec{r} \, |\varphi_\pm(\vec{r}, t)|^2 \\ \langle \vec{p}_\pm \rangle(t) &= \int \mathrm{d}^3 r \, \varphi_\pm^*(\vec{r}, t) \left[-\mathrm{i}\hbar \vec{\nabla} \varphi_\pm(\vec{r}, t)\right] \end{aligned}$$

Écrire les équations d'évolution de ces valeurs moyennes en calculant $\mathrm{d}\langle \vec{r}_\pm \rangle(t)/\mathrm{d}t$ et $\mathrm{d}\langle \vec{p}_\pm \rangle(t)/\mathrm{d}t$ à l'aide du théorème d'Eherenfest (4.26). Montrer que l'écart δ entre les centres des deux paquets d'ondes est le même que celui calculé à la question **2** pour les trajectoires classiques.

4. *Invariance par parité.* Dans une configuration expérimentale où l'on analyse un spin dirigé suivant Oz à l'aide d'un appareil de Stern-Gerlach tel que $\vec{B} \parallel Ox$, on suppose que le spin est dévié préférentiellement dans la direction $x > 0$; par exemple $\langle S_x \rangle > 0$. En examinant l'image de l'expérience dans un miroir situé dans le plan xOy, montrer qu'une telle déviation préférentielle est exclue si les interactions pertinentes dans l'expérience sont invariantes par parité (ce qui est effectivement le cas).

8.6.13 Modèle de mesure de von Neumann

1. Dans le modèle de mesure quantique imaginé par von Neumann, on se propose de mesurer une propriété physique A d'un système quantique S en faisant interagir ce système avec une particule (quantique) Π dont l'opérateur impulsion est P ; on se place pour simplifier à une dimension d'espace. Le hamiltonien d'interaction est supposé de la forme

$$H = g(t)AP$$

où $g(t)$ est une fonction positive présentant un pic aigu de largeur τ autour de $t = 0$ et telle que

$$g = \int_{-\infty}^{\infty} g(t)\mathrm{d}t \simeq \int_{-\tau/2}^{\tau/2} g(t)\mathrm{d}t$$

On suppose que l'on peut négliger l'évolution de S et de Π pendant la durée très courte τ de l'interaction entre S et Π, qui a lieu entre les temps t_i et t_f : $t_i \simeq -\tau/2$ et $t_f \simeq \tau/2$. En déduire l'opérateur d'évolution (4.14)

$$U(t_f, t_i) \simeq \mathrm{e}^{-\mathrm{i}gAP/\hbar}$$

2. On suppose que l'état initial $S + \Pi$ est

$$|\psi(t_i)\rangle = |n \otimes \varphi\rangle$$

où $|n\rangle$ est un vecteur propre de A, dont le spectre est supposé pour simplifier non dégénéré, $A|n\rangle = a_n|n\rangle$, et $|\varphi\rangle$ un état de la particule localisé au voisinage d'un point x avec une dispersion Δx. Montrer que l'état final est

$$|\psi(t_f)\rangle = |n \otimes \varphi_n\rangle \quad \text{avec} \quad |\varphi_n\rangle = \mathrm{e}^{-\mathrm{i}gAP/\hbar}\,|\varphi\rangle$$

Soit $\varphi_n(x) = \langle x|\varphi_n\rangle$, la fonction d'onde finale de la particule. Montrer que

$$\varphi_n(x) = \varphi(x - ga_n)$$

La fonction $\varphi_n(x)$ est donc localisée au voisinage du point $x - ga_n$, et si $g|a_n - a_m| \gg \Delta x$ quel que soit $n \neq m$, la position de la particule permet de remonter à la valeur a_n de A et effectue donc une mesure de A. L'état final de la particule est parfaitement corrélé à la valeur de A et à l'état final de S car les états $|\varphi_n\rangle$ et $|\varphi_m\rangle$ sont orthogonaux pour $n \neq m$: $\langle\varphi_n|\varphi_m\rangle = \delta_{nm}$.

3. Quel est l'état final de Π si l'état initial de S est la superposition linéaire

$$|\chi\rangle = \sum_n c_n|n\rangle$$

Montrer que la probabilité d'observer S dans l'état final $|n\rangle$ est $|c_n|^2$: la mesure est idéale car elle ne modifie pas les probabilités $|c_n|^2$.

8.6.14 Transformation de Galilée

Considérons une onde plane classique, par exemple une vibration sonore, se propageant suivant l'axe des x

$$f(x,t) = A\cos(kx - \omega t)$$

et une transformation de Galilée de vitesse v

$$x' = x + vt \qquad t' = t$$

1. Montrer que pour une onde classique l'amplitude transformée $f'(x',t')$ vérifie

$$f'(x',t') = f(x,t)$$

d'où l'on tire la loi de transformation des vecteurs d'onde et des fréquences

$$k' = k \qquad \omega' = \omega + vk$$

Quelle est l'interprétation physique de la loi de transformation des fréquences ? Supposons maintenant que l'on ait affaire à une onde de de Broglie pour une

particule de masse m. Les relations précédentes sont-elles compatibles avec les lois de transformation de l'impulsion et de l'énergie

$$p' = p + mv \qquad E' = E + pv + \frac{1}{2}mv^2 \ ?$$

2. Montrer que pour une onde de De Broglie on ne doit pas exiger

$$\varphi'(x', t') = \varphi(x, t)$$

mais seulement

$$\varphi'(x', t') = \exp\left[\frac{\mathrm{i}f(x,t)}{\hbar}\right]\varphi(x, t)$$

En utilisant les relations (à démontrer)

$$\begin{aligned} \frac{\partial}{\partial t'} &= \frac{\partial}{\partial t} - v\frac{\partial}{\partial x} \\ \frac{\partial}{\partial x'} &= \frac{\partial}{\partial x} \end{aligned}$$

déterminer la forme de la fonction $f(x, t)$ en exigeant que si $\varphi(x, t)$ obéit à l'équation de Schrödinger, il en soit de même pour $\varphi'(x', t')$.

8.7 Bibliographie

Les résultats de ce chapitre sont très classiques et se retrouvent sous une forme voisine dans la plupart des traités de mécanique quantique. Parmi les exposés les plus clairs, on peut retenir celui de Merzbacher [1970], chapitre 6. Lévy-Leblond et Balibar [1984], chapitre 6, donnent également une discussion très complète et illustrée par de nombreux exemples. Voir aussi Messiah [1959], chapitre III, Cohen-Tannoudji *et al.* [1973], chapitre I ou Basdevant et Dalibard [2001], chapitre 2 ; cette dernière référence est accompagnée d'un CD réalisé par M. Joffre, qui permet de visualiser le mouvement de paquets d'ondes. Pour la règle d'or de Fermi, on pourra consulter Messiah [1959], chapitre XVII ou Cohen-Tannoudji *et al.* [1973], chapitre XIII.

Chapitre 9

Moment angulaire

Dans ce chapitre, nous allons développer l'étude des propriétés du moment angulaire que nous avons introduit au chapitre 7. La propriété *fondamentale* du moment angulaire est d'être le générateur infinitésimal des rotations. Tous les résultats que nous allons obtenir dans ce chapitre sont des conséquences plus ou moins directes de cette propriété. Dans la section 9.1, nous construirons explicitement une base de vecteurs propres communs à $\vec{J}^2$ et à J_z, qui sont des opérateurs hermitiens compatibles. La rotation d'un état physique, que nous avions déjà introduite au chapitre 3 pour la polarisation d'un photon et pour le spin 1/2, sera traitée dans le cas général dans la section 9.2. La section 9.3 est consacrée au moment angulaire orbital, dont l'origine est le mouvement des particules dans l'espace. La section 9.4 transpose à la mécanique quantique les résultats classiques du mouvement dans une force centrale, tandis que la section 9.5 contient des applications aux désintégrations de particules et d'états excités. Enfin, la section 9.6 développe l'addition des moments angulaires.

N. B. Dans tout ce chapitre, on utilisera un système d'unités où $\hbar = 1$.

9.1 Diagonalisation de $\vec{J}^2$ et de J_z

Nous avons établi au chapitre 7 les relations de commutation (7.31) et (7.32) entre les différentes composantes du moment angulaire, que nous récrivons ci-dessous dans un système d'unités où $\hbar = 1$ (rappelons qu'un moment angulaire a la dimension de $\hbar$, et c'est pourquoi les notations se simplifient dans un tel système d'unités)

$$\boxed{[J_x, J_y] = \mathrm{i}J_z \qquad [J_y, J_z] = \mathrm{i}J_x \qquad [J_z, J_x] = \mathrm{i}J_y} \tag{9.1}$$

ou

$$\boxed{[J_k, J_l] = \mathrm{i}\sum_m \varepsilon_{klm} J_m} \tag{9.2}$$

La seule connaissance de ces relations de commutation va nous permettre de diagonaliser le moment angulaire, c'est-à-dire de trouver les vecteurs propres et les valeurs propres de combinaisons convenablement choisies de J_x, J_y et J_z. Comme ces trois opérateurs ne commutent pas entre eux, on ne peut évidemment pas les diagonaliser simultanément : les trois composantes de $\vec{J}$ sont mutuellement incompatibles. Une autre observation est que $\vec{J}^2$ est un opérateur scalaire et, d'après (7.33), un tel opérateur commute avec les trois composantes de $\vec{J}$

$$[\vec{J}^2, J_k] = 0 \tag{9.3}$$

ce que l'on peut vérifier par un calcul explicite (exercice 9.7.1). On choisit en général de diagonaliser simultanément $\vec{J}^2$ et J_z : c'est ce que l'on appelle souvent *quantifier le moment angulaire suivant Oz*. On dit aussi que l'axe Oz est *choisi comme axe de quantification du moment angulaire*. Il est commode de définir les opérateurs $J_\pm = J_\mp^\dagger$ par

$$J_\pm = J_x \pm \mathrm{i}J_y \tag{9.4}$$

On vérifie immédiatement les relations de commutation et les identités suivantes

$$\begin{aligned}
[J_z, J_\pm] &= \pm J_\pm && (9.5)\\
[J_+, J_-] &= 2J_z && (9.6)\\
\vec{J}^2 &= \frac{1}{2}(J_-J_+ + J_+J_-) + J_z^2 && (9.7)\\
J_+J_- &= \vec{J}^2 - J_z(J_z - 1) && (9.8)\\
J_-J_+ &= \vec{J}^2 - J_z(J_z + 1) && (9.9)
\end{aligned}$$

Ces relations nous seront utiles pour la diagonalisation. Soit $|jm\rangle$ un vecteur propre de $\vec{J}^2$ et de J_z, où j étiquette la valeur propre de $\vec{J}^2$ et m celle de J_z. Comme $\vec{J}^2$ est un opérateur positif, ses valeurs propres sont ≥ 0 et on les écrit sous la forme $j(j+1)$ avec $j \geq 0$; on justifiera bientôt cette écriture des valeurs propres de $\vec{J}^2$. Le nombre m est appelé *nombre quantique magnétique*. En résumé,

$$\begin{aligned}
\vec{J}^2|jm\rangle &= j(j+1)|jm\rangle && (9.10)\\
J_z|jm\rangle &= m|jm\rangle && (9.11)
\end{aligned}$$

D'après (9.5), les vecteurs $J_\pm|jm\rangle$ sont vecteurs propres de J_z avec la valeur propre $m \pm 1$

$$\begin{aligned}
J_z[J_\pm|jm\rangle] &= (J_\pm J_z \pm J_\pm)|jm\rangle = J_\pm(m \pm 1)|jm\rangle\\
&= (m \pm 1)[J_\pm|jm\rangle]
\end{aligned}$$

De même, comme $[\vec{J}^2, J_\pm] = 0$

$$\vec{J}^2[J_\pm|jm\rangle] = j(j+1)[J_\pm|jm\rangle]$$

Nous venons de montrer que les vecteurs $J_\pm|jm\rangle$ sont vecteurs propres de $\vec{J}^2$ avec la valeur propre $j(j+1)$ et de J_z avec la valeur propre $m \pm 1$. De plus, supposant $|jm\rangle$ de norme unité : $\langle jm|jm\rangle = 1$, on peut calculer la norme de $J_+|jm\rangle$ à partir de (9.9)

$$\begin{aligned} ||J_+|jm\rangle||^2 &= \langle jm|J_-J_+|jm\rangle = \langle jm|\vec{J}^2 - J_z(J_z+1)|jm\rangle \\ &= j(j+1) - m(m+1) = (j-m)(j+m+1) \geq 0 \end{aligned} \tag{9.12}$$

et celle de $J_-|jm\rangle$ à partir de (9.8)

$$\begin{aligned} ||J_-|jm\rangle||^2 &= \langle jm|J_+J_-|jm\rangle = \langle jm|\vec{J}^2 - J_z(J_z-1)|jm\rangle \\ &= j(j+1) - m(m-1) = (j+m)(j-m+1) \geq 0 \end{aligned} \tag{9.13}$$

La positivité n'est assurée simultanément pour les deux normes que si $-j \leq m \leq j$. Partant de $|jm\rangle$, par application répétée de J_+, on obtient une série de vecteurs propres communs à $\vec{J}^2$ et J_z, étiquetés par $(j, m+1)$, $(j, m+2)$, etc. Ces vecteurs propres ont une norme positive tant que $m \leq j$, mais la norme deviendrait négative pour $m > j$. Il faut donc que la série s'arrête, ce qui n'est possible que si l'un des vecteurs $(J_+)^n|jm\rangle$ s'annule pour une valeur entière de $n = n_1 + 1$ telle que $m + n_1 = j$

$$J_+[(J_+)^{n_1}|jm\rangle] = 0$$

Le même raisonnement avec J_- montre qu'il doit exister un entier n_2 tel que

$$J_-[(J_-)^{n_2}|jm\rangle] = 0$$

Des relations

$$j = m + n_1 \qquad -j = m - n_2$$

on déduit que $2j$, et donc $(2j+1)$ doit être un nombre entier, d'où le théorème de diagonalisation de $\vec{J}^2$ et de J_z.

Théorème. Les valeurs possibles de j sont entières ou demi-entières : $j = 0,\ 1/2,\ 1,\ 3/2, \ldots$ Si $|jm\rangle$ est vecteur propre commun à $\vec{J}^2$ et J_z, m prend nécessairement l'une des $(2j+1)$ valeurs

$$m = -j,\ -j+1,\ -j+2,\ \cdots\cdots,\ j-2,\ j-1,\ j \ \blacksquare$$

Lorsque j prend les valeurs $0,\ 1,\ 2,\ \ldots$, on dit que le moment angulaire est entier, et qu'il est demi-entier[1] lorsque $j = 1/2,\ 3/2,\ \ldots$ Examinons la normalisation et la phase des vecteurs $|jm\rangle$. Partant d'un vecteur $|jm\rangle$ on construit par application répétée de J_+ et de J_- une série de $(2j+1)$ vecteurs orthogonaux, sous-tendant un sous-espace vectoriel à $(2j+1)$ dimensions $\mathcal{E}(j)$

1. Bien que la moitié d'un entier pair soit aussi un demi-entier...

de $\mathcal{H}$. Ces vecteurs ne sont pas de norme unité, mais si l'on définit $|j, m-1\rangle$ par

$$|j, m-1\rangle = [j(j+1) - m(m-1)]^{-1/2} J_- |jm\rangle \tag{9.14}$$

alors $|j, m-1\rangle$ est bien de norme unité d'après (9.13). D'autre part, en utilisant (9.8)

$$\begin{aligned} J_+ J_- |jm\rangle &= [j(j+1) - m(m-1)]^{1/2} J_+ |j, m-1\rangle \\ &= [j(j+1) - m(m-1)] |jm\rangle \end{aligned}$$

soit

$$J_+ |j, m-1\rangle = [j(j+1) - m(m-1)]^{1/2} |jm\rangle$$

ou bien, avec la substitution $m \to m+1$

$$J_+ |jm\rangle = [j(j+1) - m(m+1)]^{1/2} |j, m+1\rangle \tag{9.15}$$

Les relations (9.14) ou (9.15) fixent complètement la phase relative des vecteurs $|j, j\rangle,\ |j, j-1\rangle,\ \ldots, |j, -j\rangle$. Une base de $\mathcal{E}(j)$ formée de vecteurs $|jm\rangle$ qui satisfont à (9.14) ou (9.15) est appelée *base standard* $|jm\rangle$.

Il peut arriver que la donnée de (j, m) ne suffise pas à spécifier de façon unique un vecteur de $\mathcal{H}$: $\vec{J}^{\,2}$ et J_z ne forment pas un ensemble complet de propriétés physiques compatibles. Nous en verrons un exemple au § 9.4.2 avec l'atome d'hydrogène : la donnée du moment angulaire (orbital), noté dans ce cas l, ne suffit pas à spécifier un état lié, il faut en plus se donner un nombre quantique $n = l+1,\ l+2,\ \cdots$, ou nombre quantique principal. En général, on doit utiliser un nombre quantique, ou un ensemble de nombres quantiques supplémentaires τ pour étiqueter des vecteurs propres $|j, m=j\rangle$ de $\vec{J}^{\,2}$ et de J_z, normalisés par

$$\langle \tau, j, j | \tau', j, j\rangle = \delta_{\tau,\tau'}$$

Par application répétée de J_-, on forme une base standard de $\mathcal{E}(\tau, j)$

$$|\tau, j, j\rangle,\ |\tau, j, j-1\rangle,\ \ldots\ldots, |\tau, j, -j+1\rangle,\ \ |\tau, j, -j\rangle$$

Résumons les propriétés essentielles d'une base standard $|\tau, jm\rangle$

$$\vec{J}^{\,2} |\tau, jm\rangle = j(j+1) |\tau, jm\rangle \qquad J_z |\tau, jm\rangle = m |\tau, jm\rangle \tag{9.16}$$

$$J_+ |\tau, jm\rangle = [j(j+1) - m(m+1)]^{1/2} |\tau, j, m+1\rangle \tag{9.17}$$

$$J_- |\tau, jm\rangle = [j(j+1) - m(m-1)]^{1/2} |\tau, j, m-1\rangle \tag{9.18}$$

$$J_+ |\tau, j, j\rangle = 0 \qquad J_- |\tau, j, -j\rangle = 0 \tag{9.19}$$

$$\langle \tau', j'm' | \tau, jm\rangle = \delta_{\tau'\tau} \delta_{j'j} \delta_{m'm} \tag{9.20}$$

Nous supprimerons par la suite l'indice τ qui ne jouera aucun rôle dans ce chapitre. Les éléments de matrice de $\vec{J}^2$, J_z et $J_\pm$ dans une base standard sont

$$\langle j'm'|\vec{J}^2|jm\rangle = j(j+1)\delta_{j'j}\delta_{m'm} \tag{9.21}$$

$$\langle j'm'|J_z|jm\rangle = m\,\delta_{j'j}\delta_{m'm} \tag{9.22}$$

$$\langle j'm'|J_\pm|jm\rangle = [j(j+1)-mm']^{1/2}\delta_{j'j}\delta_{m',m\pm 1} \tag{9.23}$$

Dans le sous-espace $\mathcal{E}(j)$ où $\vec{J}^2$ a une valeur propre $j(j+1)$ fixée, les opérateurs J_z et $J_\pm$ sont représentés par des matrices $(2j+1)\times(2j+1)$, la matrice représentant J_z étant diagonale. Il est instructif (exercice 9.7.4) d'expliciter ces matrices dans le cas $j = 1/2$ et de retouver les matrices 2×2 (3.47) du spin 1/2, ainsi que dans le cas $j = 1$, où l'on retrouve les générateurs infinitésimaux des rotations dans l'espace à trois dimensions : la loi de transformation d'un vecteur est celle d'un moment angulaire $j = 1$. L'équation (9.23) donne pour les générateurs infinitésimaux du cas $j = 1$ (exercice 9.7.4)

$$\boxed{J_x = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad J_y = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix} \quad J_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}} \tag{9.24}$$

Ces générateurs infinitésimaux ont superficiellement une forme différente de celle des générateurs T_i déterminés en (7.26). En fait, ils y sont reliés par la transformation unitaire (9.64) qui fait passer des composantes cartésiennes de $\hat{r}$ à ses composantes sphériques : exercice 9.7.4.

9.2 Matrices de rotation

Nous avons vu au chapitre 3 comment faire tourner un spin 1/2 : partant d'un état $|+\rangle$ obtenu à l'aide d'un appareil de Stern-Gerlach dont le champ magnétique est parallèle à Oz, nous savons construire grâce à (3.57) l'état $|+,\hat{n}\rangle$ obtenu à l'aide d'un appareil de Stern-Gerlach dont le champ magnétique est parallèle à $\hat{n}$. On fait agir sur l'état $|+\rangle$ un opérateur de rotation $U[\mathcal{R}]$ qui transforme $|+\rangle$ en $|+,\hat{n}\rangle$

$$|+,\hat{n}\rangle = U[\mathcal{R}]|+\rangle = |+\rangle_{\mathcal{R}}$$

La rotation $\mathcal{R}$ amène l'axe Oz sur la direction $\hat{n}$; cette rotation n'est pas unique, et nous verrons que cette absence d'unicité correspond à l'arbitraire de phase dans la définition de $|+,\hat{n}\rangle$. Un autre exemple de rotation d'un état physique a été donné au chapitre 3 dans le cas de la polarisation d'un photon : partant d'un état de polarisation linéaire $|x\rangle$, on obtient un état de polarisation linéaire $|\theta\rangle$ en lui appliquant un opérateur de rotation $U[\mathcal{R}_z(\theta)]$ correspondant à une rotation d'angle θ autour de la direction de propagation Oz du photon (3.29)

$$|\theta\rangle = \exp(-\mathrm{i}\theta\Sigma_z)|x\rangle = U[\mathcal{R}_z(\theta)]\,|x\rangle$$

Dans le cas général, l'état $|\varphi\rangle_{\mathcal{R}}$ transformé par une rotation $\mathcal{R}$ d'un état $|\varphi\rangle$ est

$$|\varphi\rangle_{\mathcal{R}} = U[\mathcal{R}]\,|\varphi\rangle$$

Nous allons donner la forme matricielle explicite des opérateurs de rotation $U[\mathcal{R}]$ dans la base $|jm\rangle$. L'opérateur de rotation $U[\mathcal{R}]$ s'exprime en fonction des générateurs infinitésimaux J_x, J_y et J_z : *cf.* (7.30). Comme les composantes de $\vec{J}$ commutent avec $\vec{J}^2$, le commutateur $[U(\mathcal{R}), \vec{J}^2] = 0$ et les éléments de matrice de U sont nuls si $j \neq j'$

$$\langle j'm'|U[\mathcal{R}]|jm\rangle \propto \delta_{j'j}$$

Dans le sous-espace $\mathcal{E}(j)$, l'opérateur $U(\mathcal{R})$ sera représenté par une matrice $(2j+1) \times (2j+1)$ notée $D^{(j)}[\mathcal{R}]$. Ses éléments de matrice sont

$$D^{(j)}_{m'm}[\mathcal{R}] = \langle jm'|U[\mathcal{R}]|jm\rangle \tag{9.25}$$

Les matrices $D^{(j)}$ sont appelées *matrices de rotation*, ou *matrices de Wigner.* Examinons la transformation par rotation d'un état $|jm\rangle$ qui donne le vecteur $|jm\rangle_{\mathcal{R}}$

$$|jm\rangle_{\mathcal{R}} = U[\mathcal{R}]\,|jm\rangle = \sum_{m'} |jm'\rangle\langle jm'|\,U[\mathcal{R}]\,|jm\rangle$$

où l'on a remarqué que dans la relation de fermeture

$$\sum_{j',m'} |j'm'\rangle\langle j'm'| = I$$

seuls les termes $j = j'$ contribuent. On peut donc écrire

$$\boxed{|jm\rangle_{\mathcal{R}} = \sum_{m'} D^{(j)}_{m'm}[\mathcal{R}]|jm'\rangle} \tag{9.26}$$

Rappelons les propriétés de groupe des opérateurs $U[\mathcal{R}]$: dans le cas d'une représentation vectorielle (7.12)

$$U[\mathcal{R}_2]U[\mathcal{R}_1] = U[\mathcal{R}_2\mathcal{R}_1] \tag{9.27}$$

et dans celui d'une représentation spinorielle (7.13)

$$U[\mathcal{R}_2]U[\mathcal{R}_1] = \pm U[\mathcal{R}_2\mathcal{R}_1] \tag{9.28}$$

Nous montrerons à la fin de ce paragraphe que (9.27) correspond au cas d'un moment angulaire entier et (9.28) à celui d'un moment angulaire demi-entier. À la propriété de groupe pour les opérateurs U correspond une propriété analogue pour les matrices de rotation

$$D^{(j)}_{m'm}[\mathcal{R}_2\mathcal{R}_1] = \pm \sum_{m''} D^{(j)}_{m'm''}[\mathcal{R}_2]D^{(j)}_{m''m}[\mathcal{R}_1]$$

Revenons à l'étude de la rotation qui amène l'axe Oz dans la direction $\hat{n}$ d'angles polaire et azimutal (θ, ϕ)

$$\hat{n}_x = \sin\theta\cos\phi \qquad \hat{n}_y = \sin\theta\sin\phi \qquad \hat{n}_z = \cos\theta \tag{9.29}$$

Nous allons choisir *par convention* la rotation suivante : $\mathcal{R}$, noté $\mathcal{R}(\theta, \phi)$, sera le produit d'une première rotation d'angle θ autour de Oy suivie d'une rotation de ϕ autour de Oz (figure 9.1)

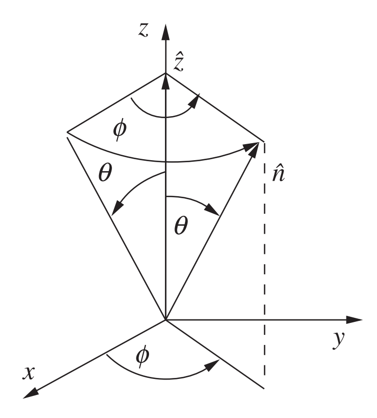


Fig. 9.1 – Rotation $\mathcal{R}(\theta, \phi)$ amenant l'axe Oz sur $\hat{n}$.

$$\boxed{\mathcal{R}(\theta, \phi) = \mathcal{R}_z(\phi)\mathcal{R}_y(\theta)} \tag{9.30}$$

Compte tenu de (9.30) et de la loi de groupe, l'opérateur de rotation $U[\mathcal{R}(\theta, \phi)]$ est donné en fonction des générateurs infinitésimaux J_y et J_z par

$$U[\mathcal{R}(\theta, \phi)] = \mathrm{e}^{-\mathrm{i}\phi J_z}\,\mathrm{e}^{-\mathrm{i}\theta J_y} \tag{9.31}$$

et ses éléments de matrice dans une base $|jm\rangle$ par

$$\boxed{D^{(j)}_{m'm}[\mathcal{R}(\theta, \phi)] = \langle jm'|\mathrm{e}^{-\mathrm{i}\phi J_z}\,\mathrm{e}^{-\mathrm{i}\theta J_y}|jm\rangle} \tag{9.32}$$

Cette équation se simplifie en

$$D^{(j)}_{m'm}[\mathcal{R}(\theta, \phi)] \equiv D^{(j)}_{m'm}(\theta, \phi) = \mathrm{e}^{-\mathrm{i}m'\phi}\langle jm'|\mathrm{e}^{-\mathrm{i}\theta J_y}|jm\rangle \tag{9.33}$$

$$= \mathrm{e}^{-\mathrm{i}m'\phi}\, d^{(j)}_{m'm}(\theta) \tag{9.34}$$

Nous avons défini la matrice $d^{(j)}(\theta)$ par

$$\boxed{d^{(j)}_{m'm}(\theta) = \langle jm'|\mathrm{e}^{-\mathrm{i}\theta J_y}|jm\rangle} \tag{9.35}$$

Les matrices $d^{(j)}$ vérifient une propriété de groupe déduite de celle des matrices $D^{(j)}$

$$d^{(j)}_{m'm}(\theta_2+\theta_1)=\sum_{m''} d^{(j)}_{m'm''}(\theta_2)d^{(j)}_{m''m}(\theta_1)$$

Il n'y pas de signe $\pm$ dans cette équation car l'angle de rotation peut être supérieur à 2π.

Nous avons déjà mentionné l'arbitraire dans le choix de la rotation (θ,ϕ) : nous aurions pu effectuer une première rotation d'un angle ψ autour de Oz sans modifier l'axe $\hat{n}$ final. Le nouvel opérateur de rotation serait

$$U[\mathcal{R}']=U[\mathcal{R}(\theta,\phi)]\mathrm{e}^{-\mathrm{i}\psi J_z}$$

et le résultat (9.26) serait modifié par le facteur de phase $\exp(-\mathrm{i}m\psi)$. La définition la plus générale des matrices de rotation fait intervenir trois angles, appelés angles d'Euler, (ϕ,θ,ψ), et notre convention correspond au choix d'angles d'Euler[2] $(\phi,\theta,0)$.

Dans la base $|jm\rangle$, $\mathrm{i}J_y$ est représenté par une matrice réelle, car d'après (9.23) les éléments de matrice de J_+ et de J_- sont réels et

$$J_y=-\frac{\mathrm{i}}{2}\,(J_+-J_-)$$

La matrice $\exp(-\mathrm{i}\theta J_y)$ est aussi une matrice réelle et la propriété de groupe

$$U^\dagger[\mathcal{R}]=U^{-1}[\mathcal{R}]=U[\mathcal{R}^{-1}]$$

se traduit par

$$\left[d^{(j)}(\theta)\right]^\dagger=\left[d^{(j)}(-\theta)\right]$$

ce qui donne pour les éléments de matrice

$$\boxed{d^{(j)}_{m'm}(\theta)=d^{(j)}_{mm'}(-\theta)} \qquad (9.36)$$

Il existe une autre propriété de symétrie (exercice 9.7.12)

$$\boxed{d^{(j)}_{m'm}(\theta)=(-1)^{m-m'}\,d^{(j)}_{-m',-m}(\theta)} \qquad (9.37)$$

Enfin, on peut montrer que les matrices $D^{(j)}$ forment une représentation dite *irréductible* du groupe des rotations, c'est-à-dire que tout vecteur de $\mathcal{E}(j)$ peut être obtenu à partir d'un vecteur arbitraire de cet espace par application

2. La notation habituelle pour les matrices de rotation est

$$D^{(j)}(\theta,\phi)\to D^{(j)}(\phi,\theta,\psi=0)$$

d'une matrice de rotation $D^{(j)}$, et que toute matrice commutant avec toutes les matrices $D^{(j)}$ est multiple de la matrice identité.

Le test pour la présence du facteur $\pm$ dans (9.28) est donné par l'examen des rotations de 2π : ce facteur apparaît lorsqu'une rotation de 2π est représentée par l'opérateur $-I$ dans l'espace $\mathcal{E}(j)$. Examinons une rotation de 2π autour de l'axe Oz

$$\begin{aligned}\langle jm'|U[\mathcal{R}_z(2\pi)]|jm\rangle &= \mathrm{e}^{-2\mathrm{i}\pi m}\,\delta_{m'm} = \delta_{m'm} \quad \text{j entier}\\ &= \mathrm{e}^{-2\mathrm{i}\pi m}\,\delta_{m'm} = -\delta_{m'm} \quad \text{j demi-entier}\end{aligned}$$

Comme le choix de l'axe Oz est arbitraire, l'opérateur de rotation d'angle 2π vaut I pour j entier et $-I$ pour j demi-entier. En revanche, les opérateurs de rotation d'angle 4π sont égaux à I pour toute valeur de j. Examinons deux rotations successives d'angles θ_1 et θ_2 autour d'un axe $\hat{n}$, avec

$$\theta_1 + \theta_2 = \theta + 2\pi n \qquad 0 \le \theta < 2\pi \qquad n \text{ entier} \ge 0$$

Des équations

$$\begin{aligned}\mathrm{e}^{-\mathrm{i}(\theta_1+\theta_2)\vec{J}\cdot\hat{n}} = \mathrm{e}^{-\mathrm{i}\theta(\vec{J}\cdot\hat{n})}\mathrm{e}^{-2\mathrm{i}\pi n(\vec{J}\cdot\hat{n})} &= \mathrm{e}^{-\mathrm{i}\theta(\vec{J}\cdot\hat{n})} \quad j \text{ entier}\\ &= (-1)^n\,\mathrm{e}^{-\mathrm{i}\theta(\vec{J}\cdot\hat{n})} \quad j \text{ demi-entier}\end{aligned}$$

on déduit que (9.27) est valable pour j entier et (9.28) pour j demi-entier. En d'autres termes, à toute rotation $\mathcal{R}$ correspondent deux opérateurs de rotation de signe opposé pour j demi-entier, et un seul pour j entier.

Dans le cas du spin 1/2, vérifions que nous retrouvons bien la matrice $D^{(1/2)}(\theta,\phi)$ déjà calculée au chapitre 3. La matrice $d^{(1/2)}(\theta)$ vaut d'après (3.59)

$$d^{(1/2)}(\theta) = \exp(-\mathrm{i}\theta\sigma_y/2) = \cos\frac{\theta}{2}\,I - \mathrm{i}\sigma_y\sin\frac{\theta}{2}$$

ou sous forme explicite

$$d^{(1/2)}(\theta) = \begin{pmatrix} \cos\theta/2 & -\sin\theta/2 \\ \sin\theta/2 & \cos\theta/2 \end{pmatrix} \tag{9.38}$$

les lignes et les colonnes étant rangées dans l'ordre $m = 1/2, -1/2$; l'équation (9.33) donne alors la matrice $D^{(j)}(\theta,\phi)$

$$D^{(1/2)}(\theta,\phi) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\phi/2}\cos\theta/2 & -\mathrm{e}^{-\mathrm{i}\phi/2}\sin\theta/2 \\ \mathrm{e}^{\mathrm{i}\phi/2}\sin\theta/2 & \mathrm{e}^{\mathrm{i}\phi/2}\cos\theta/2 \end{pmatrix}$$

en accord avec (3.58).

La matrice de rotation $d^{(1)}(\theta)$ d'un moment angulaire $j = 1$ s'obtient à partir de l'expression (9.24) des générateurs infinitésimaux pour le moment

angulaire un, les lignes et les colonnes étant rangées dans l'ordre $m = 1, 0, -1$ (exercice 9.7.4)

$$d^{(1)}(\theta) = \begin{pmatrix} \frac{1}{2}(1+\cos\theta) & -\frac{1}{\sqrt{2}}\sin\theta & \frac{1}{2}(1-\cos\theta) \\ \frac{1}{\sqrt{2}}\sin\theta & \cos\theta & -\frac{1}{\sqrt{2}}\sin\theta \\ \frac{1}{2}(1-\cos\theta) & \frac{1}{\sqrt{2}}\sin\theta & \frac{1}{2}(1+\cos\theta) \end{pmatrix} \tag{9.39}$$

On vérifie que les matrices $d^{(1/2)}$ et $d^{(1)}$ obéissent aux propriétés de symétrie (9.36) et (9.37).

9.3 Moment angulaire orbital

9.3.1 Opérateur moment angulaire orbital

Considérons un champ scalaire classique $\psi(\vec{r})$ et faisons, lui subir une rotation $R_z(\phi)$ d'angle ϕ autour de Oz ; soit $\vec{r}\,' = \mathcal{R}\vec{r}$ le vecteur transformé du vecteur $\vec{r}$ dans cette rotation

$$\begin{aligned} x' &= x\cos\phi - y\sin\phi \\ y' &= x\sin\phi + y\cos\phi \\ z' &= z \end{aligned}$$

La valeur du champ scalaire transformé $\psi'(\vec{r})$ au point $\vec{r}\,'$ doit être identique à sa valeur au point $\vec{r}$

$$\psi'(\vec{r}\,') = \psi(\vec{r})$$

ou encore

$$\boxed{\psi'(\vec{r}) = \psi(\mathcal{R}^{-1}\vec{r})} \tag{9.40}$$

Cette loi de transformation est correcte pour un champ scalaire classique, mais si $\psi(\vec{r})$ est une fonction d'onde, $\psi(\mathcal{R}^{-1}\vec{r})$ et $\psi'(\vec{r})$ pourraient *a priori* différer par un facteur de phase

$$\psi'(\vec{r}) = \mathrm{e}^{\mathrm{i}\theta(\vec{r})}\,\psi(\mathcal{R}^{-1}\vec{r})$$

(voir la discussion suivant (8.17)). En toute rigueur, nous savons seulement que $|\psi'(\vec{r}\,')| = |\psi(\vec{r})|$ et nous devons montrer que le facteur de phase éventuel est absent. Le vecteur $U(\mathcal{R})|\vec{r}\rangle$ représente physiquement un état propre de l'opérateur position $\vec{R}$, obtenu à partir de l'état propre $|\vec{r}\rangle$ de $\vec{R}$ par une rotation $U(\mathcal{R})$. Vérifions-le explicitement en utilisant le fait que $\vec{R}$ est un opérateur vectoriel dont les composantes X_k se transforment suivant (8.34)

$$\begin{aligned} X_k[U(\mathcal{R})|\vec{r}\rangle] &= U(\mathcal{R})U^{-1}(\mathcal{R})X_k U(\mathcal{R})|\vec{r}\rangle \\ &= U(\mathcal{R})\Big(\sum_l \mathcal{R}_{kl}X_l\Big)|\vec{r}\rangle = U(\mathcal{R})\Big(\sum_l \mathcal{R}_{kl}x_l\Big)|\vec{r}\rangle \\ &= (\mathcal{R}\vec{r}\,)_k\,[U(\mathcal{R})|\vec{r}\rangle] \end{aligned}$$

ce qui montre que l'on peut définir le vecteur d'état $|\mathcal{R}\vec{r}\,\rangle$, c'est-à-dire fixer sa phase, par

$$|\mathcal{R}\vec{r}\,\rangle \equiv U(\mathcal{R})|\vec{r}\,\rangle \tag{9.41}$$

Si $|\psi'\rangle$ est le transformé par $U(\mathcal{R})$ de $|\psi\rangle$: $|\psi'\rangle = U(\mathcal{R})|\psi\rangle$, alors

$$\begin{aligned}\psi'(\vec{r}\,) &= \langle\vec{r}\,|\psi'\rangle = \langle\vec{r}\,|U(\mathcal{R})|\psi\rangle = \langle U^\dagger(\mathcal{R})\vec{r}\,|\psi\rangle \\ &= \langle U^{-1}(\mathcal{R})\vec{r}\,|\psi\rangle = \langle\mathcal{R}^{-1}\vec{r}\,|\psi\rangle = \psi(\mathcal{R}^{-1}\vec{r})\end{aligned}$$

ce qui démontre (9.40). À première vue, l'argument $\mathcal{R}^{-1}$ dans (9.40) qui s'écrit aussi

$$[U(\mathcal{R})\psi](\vec{r}) = \psi(\mathcal{R}^{-1}\vec{r})$$

peut surprendre, mais nous avons déjà rencontré une situation analogue pour les translations dans l'équation (8.15), écrite ci-dessous dans le cas de la dimension 3 avec $\hbar = 1$

$$\left[\mathrm{e}^{-\mathrm{i}\vec{P}\cdot\vec{a}}\,\psi\right](\vec{r}) = \psi(\vec{r} - \vec{a})$$

alors que[3]

$$\mathrm{e}^{-\mathrm{i}\vec{P}\cdot\vec{a}}\,|\vec{r}\,\rangle = |\vec{r} + \vec{a}\,\rangle$$

La transformée par translation de $\vec{a}$ de la fonction $\psi(\vec{r})$ est bien $\psi(\vec{r} - \vec{a})$ et non $\psi(\vec{r} + \vec{a})$! Si l'angle de rotation ϕ devient infinitésimal pour une rotation autour de Oz

$$U[\mathcal{R}_z(\phi)] \simeq I - \mathrm{i}\phi J_z$$

et d'après (9.40)

$$\begin{aligned}[(I - \mathrm{i}\phi J_z)\psi](\vec{r}) &\simeq \psi(x + y\phi, -x\phi + y, z) \\ &\simeq \psi(\vec{r}) + \phi\left(y\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial y}\right) \\ &= \psi(\vec{r}) - \mathrm{i}\phi(XP_y - YP_x)\psi\end{aligned}$$

d'où

$$[J_z\psi](\vec{r}) = [(XP_y - YP_x)\psi](\vec{r}) = [(\vec{R} \times \vec{P})_z\psi](\vec{r}) = (\vec{L}_z\psi)(\vec{r}) \tag{9.42}$$

L'opérateur moment angulaire de la particule décrite par une fonction d'onde $\psi(\vec{r})$ est appelé *moment angulaire orbital* (car associé au mouvement d'une particule sur une orbite dans l'espace), et il est en général noté $\vec{L}$

$$\vec{L} = \vec{R} \times \vec{P} \tag{9.43}$$

$\vec{L}$ a été construit comme générateur infinitésimal de rotations et vérifie nécessairement les relations de commutation d'un moment angulaire (9.1) ou (9.2)

$$[L_j, L_k] = \mathrm{i}\sum_l \varepsilon_{jkl}L_l \tag{9.44}$$

3. On remarquera que cette équation fixe la phase du vecteur $|\vec{r} + \vec{a}\,\rangle$ relativement à celle de $|\vec{r}\,\rangle$, de même que (9.41) fixe la phase de $|\mathcal{R}\vec{r}\,\rangle$ relativement à celle de $|\vec{r}\,\rangle$.

On peut vérifier ces relations par un calcul explicite utilisant les relations de commutation canoniques (7.45) : exercice 9.7.5. On note $|lm\rangle$ les vecteurs propres de $\vec{L}^{\,2}$ et de L_z

$$\vec{L}^{\,2}|lm\rangle = l(l+1)|lm\rangle \tag{9.45}$$

$$L_z|lm\rangle = m|lm\rangle \tag{9.46}$$

Ces équations peuvent être transformées en équations différentielles, si l'on écrit les opérateurs L_j sous forme d'opérateurs différentiels agissant dans $L^{(2)}(\mathbb{R}^3)$. Le calcul est laborieux si l'on effectue le changement de variables $(x, y, z) \to (r, \theta, \phi)$ mais il se simplifie si l'on utilise la propriété des L_i d'être les générateurs infinitésimaux des rotations. Le cas de L_z est particulièrement simple : considérant ψ comme fonction de (r, θ, ϕ), nous avons

$$\left[\mathrm{e}^{\,-\mathrm{i}\alpha L_z}\,\psi\right](r,\theta,\phi) = \psi(r,\theta,\phi-\alpha)$$

et en prenant α infinitésimal

$$\left[(I - \mathrm{i}\alpha L_z)\psi\right](r,\theta,\phi) = \psi(r,\theta,\phi) - \alpha\frac{\partial\psi}{\partial\phi}$$

soit $L_z\psi = -\mathrm{i}(\partial\psi/\partial\phi)$. Le calcul de L_x et L_y prend quelques lignes de plus, car dans une rotation autour de Ox ou autour de Oy, les angles θ et ϕ varient tous les deux ; on trouve (exercice 9.7.5)

$$L_z = -\mathrm{i}\frac{\partial}{\partial\phi} \tag{9.47}$$

$$L_\pm = \mathrm{i}\,\mathrm{e}^{\,\pm\mathrm{i}\phi}\left(\cot\theta\frac{\partial}{\partial\phi} \mp \mathrm{i}\frac{\partial}{\partial\theta}\right) \tag{9.48}$$

$$\vec{L}^{\,2} = -\left(\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\left[\sin\theta\frac{\partial}{\partial\theta}\right] + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\phi^2}\right) \tag{9.49}$$

Les opérateurs L_j dépendent seulement des *angles*, et non de r, d'où la dénomination *moment angulaire*. Les fonctions propres de $\vec{L}^{\,2}$ et de L_z ne dépendent que des angles θ et ϕ ou, de façon équivalente, de $\hat{r}$. Ces fonctions propres sont appelées *harmoniques sphériques*

$$Y_l^m(\theta,\phi) = Y_l^m(\hat{r}) = \langle\hat{r}|lm\rangle \tag{9.50}$$

Les équations (9.45) et (9.46) deviennent

$$[\vec{L}^{\,2}Y_l^m](\hat{r}) = \langle\hat{r}\,|\vec{L}^{\,2}|lm\rangle = l(l+1)Y_l^m(\hat{r}) \tag{9.51}$$

$$[L_zY_l^m](\hat{r}) = \langle\hat{r}\,|L_z|lm\rangle = mY_l^m(\hat{r}) \tag{9.52}$$

tandis que (9.15) s'écrit

$$[L_\pm Y_l^m](\hat{r}) = \langle\hat{r}\,|L_\pm|lm\rangle = [l(l+1) - m(m+1)]^{1/2}\,Y_l^{m\pm1}(\hat{r})$$

L'equation (9.52) devient, compte tenu de (9.47)

$$[L_z Y_l^m](\theta,\phi) = -\mathrm{i}\frac{\partial}{\partial\phi}\, Y_l^m(\theta,\phi) = mY_l^m(\theta,\phi)$$

ce qui implique

$$Y_l^m(\theta,\phi) = \mathrm{e}^{\,\mathrm{i}m\phi}\, f_l^m(\theta) \tag{9.53}$$

La loi de tranformation (9.40) montre que dans une rotation de 2π la fonction d'onde est inchangée, et qu'il ne s'introduit donc aucun signe $(-)$ dans une telle rotation. Ceci implique que *les moments angulaires orbitaux sont toujours entiers.*

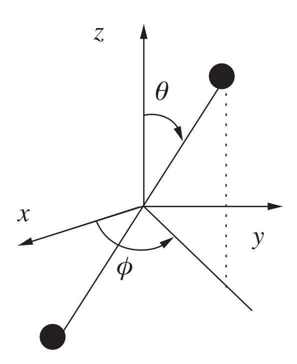


FIG. 9.2 – Rotateur sphérique.

Une application simple et importante est le rotateur sphérique : considérons une molécule diatomique en rotation autour de son centre de masse pris comme origine des coordonnées (figure 9.2 et exercice 1.6.1). Son moment d'inertie est $I = \mu r_0^2$, où μ est la masse réduite et r_0 la distance entre les noyaux (la contribution des électrons est négligeable). Si ω est la vitesse angulaire de rotation, le hamiltonien classique H_{cl} est

$$H_{\mathrm{cl}} = \frac{1}{2}\, I\omega^2 = \frac{1}{2}\,\frac{(I\omega)^2}{I} = \frac{l^2}{2I}$$

où $l = I\omega$ est le moment angulaire. La version quantique du hamiltonien est

$$H = \frac{\vec{L}^2}{2I}$$

et les valeurs de l'énergie sont

$$E_l = \frac{l(l+1)}{2I} \tag{9.54}$$

Les fonctions propres sont les $Y_l^m(\theta,\phi)$, où les angles θ et ϕ donnent l'orientation de la droite joignant les deux noyaux : $Y_l^m(\theta,\phi)$ est l'amplitude pour

trouver cette droite orientée suivant la direction (θ, ϕ). Le spectre des niveaux de rotation est donné sur la figure 9.3, et il reproduit bien les données expérimentales sur les spectres des molécules diatomiques. Une justification plus complète de (9.54) sera donnée au § 16.3.2.

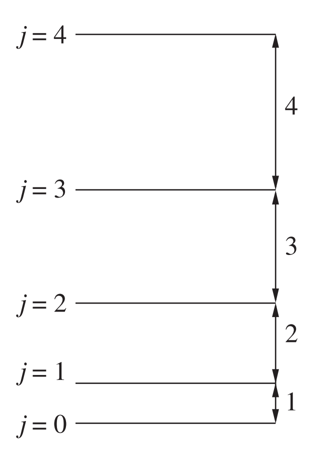


FIG. 9.3 – Spectre du rotateur sphérique. Le niveau j est séparé du niveau $(j-1)$ par j/I, ou $\hbar^2 j/I$ si l'on rétablit $\hbar$.

9.3.2 Propriétés des harmoniques sphériques

Nous allons maintenant résumer les principales propriétés des harmoniques sphériques, celles qui sont les plus souvent utilisées.

1. *Base sur la sphère unité.* Les harmoniques sphériques forment une base orthonormée pour les fonctions de carré sommable sur la sphère unité $\vec{r}^{\,2} = 1$

$$\int \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi \, [Y_{l'}^{m'}(\theta,\phi)]^* \, Y_l^m(\theta,\phi) = \int \mathrm{d}\Omega \, [Y_{l'}^{m'}(\theta,\phi)]^* \, Y_l^m(\theta,\phi) = \delta_{l'l}\delta_{m'm} \tag{9.55}$$

On utilise fréquemment la notation angle solide : $\Omega = (\theta, \phi)$ et

$$\mathrm{d}\Omega = \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi = \mathrm{d}^2\hat{r} \tag{9.56}$$

Si une fonction $f(\theta, \phi)$ est de carré sommable sur la sphère unité, on peut écrire un développement analogue à un développement de Fourier

$$\begin{aligned} f(\theta,\phi) &= \sum_{l.m} c_{lm} Y_l^m(\theta,\phi) \\ c_{lm} &= \int \mathrm{d}\Omega \, [Y_l^m(\theta,\phi)]^* \, f(\theta,\phi) \end{aligned} \tag{9.57}$$

2. *Relation avec les polynômes de Legendre.* Une définition possible des polynômes de Legendre $P_l(u)$ est

$$P_l(u) = \frac{1}{2^l\, l!}\,\frac{\mathrm{d}^l}{\mathrm{d}u^l}(u^2-1)^l \qquad (9.58)$$

$P_l(u)$ est un polynôme de degré l, de parité $(-1)^l$

$$P_l(-u) = (-1)^l P_l(u)$$

Les polynômes de Legendre forment un système complet de polynômes orthogonaux dans l'intervalle $[-1,+1]$. Les premiers polynômes de Legendre sont

$$P_0(u) = 1 \qquad P_1(u) = u \qquad P_2(u) = \frac{1}{2}\,(3u^2-1) \qquad (9.59)$$

Les fonctions de Legendre associées $P_l^m(u)$ sont définies par

$$P_l^m(u) = (1-u^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}u^m}\,P_l(u) \qquad P_l^0(u) = P_l(u) \qquad (9.60)$$

et on montre que les harmoniques sphériques sont reliés aux P_l^m par

$$\begin{aligned} Y_l^m(\theta,\phi) &= (-1)^m\left[\frac{(2l+1)}{4\pi}\,\frac{(l-m)!}{(l+m)!}\right]^{1/2} P_l^m(\cos\theta)\,\mathrm{e}^{\mathrm{i}m\phi} \;&:\; m>0 \\ Y_l^m(\theta,\phi) &= (-1)^m[Y_l^{-m}(\theta,\phi)]^* \;&:\; m<0 \end{aligned} \qquad (9.61)$$

D'après (9.53), Y_l^0 est indépendant de ϕ et est proportionnel à $P_l(\cos\theta)$

$$Y_l^0(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}}\,P_l(\cos\theta) \qquad (9.62)$$

Comme cas particuliers, écrivons Y_l^m pour $l=0$ et $l=1$

$$\begin{aligned} l=0 \quad & Y_0^0 = \sqrt{\frac{1}{4\pi}} \\ l=1 \quad & Y_1^0 = \sqrt{\frac{3}{4\pi}}\,\cos\theta = \sqrt{\frac{3}{4\pi}}\,\hat{r}_0 \\ & Y_1^{\pm} = \mp\sqrt{\frac{3}{8\pi}}\,\mathrm{e}^{\pm\mathrm{i}\phi}\,\sin\theta = \sqrt{\frac{3}{4\pi}}\,\hat{r}_{\pm1} \end{aligned} \qquad (9.63)$$

À un facteur de normalisation $\sqrt{3/4\pi}$ près, les Y_1^m ne sont pas autre chose que les *composantes sphériques* du vecteur unitaire $\hat{r}$

$$\hat{r} = (\sin\theta\cos\phi,\, \sin\theta\sin\phi,\, \cos\theta)$$

$$Y_1^0 = \sqrt{\frac{3}{4\pi}}\,\hat{r}_0 \qquad Y_1^{\pm} = \mp\sqrt{\frac{3}{4\pi}}\,\frac{\hat{r}_x \pm \mathrm{i}\hat{r}_y}{\sqrt{2}} = \sqrt{\frac{3}{4\pi}}\,\hat{r}_{\pm1} \qquad (9.64)$$

Ces formules justifient les conventions de phase utilisées pour les polarisations droite et gauche dans (3.11).
3. *Transformation par rotation.* En multipliant à gauche la relation (9.26) écrite pour $j = l$ par le bra $\langle \hat{r}\,|$, on obtient

$$Y_l^m(\mathcal{R}^{-1}\hat{r}) = \sum_{m'} D^{(l)}_{m'm}(\mathcal{R})\, Y_l^{m'}(\hat{r}) \tag{9.65}$$

On peut montrer également (exercice 9.7.6) une relation entre harmoniques sphériques et matrices de rotation

$$D^{(l)}_{m0}(\theta,\phi) = \sqrt{\frac{4\pi}{2l+1}}\,[Y_l^m(\theta,\phi)]^* \tag{9.66}$$

On déduit de ces deux équations le théorème d'addition des harmoniques sphériques. En effet, prenant $\hat{r}$ suivant la direction d'angles polaires (α,β), soit $\mathcal{R}$ la rotation d'angles (θ,ϕ) amenant $\hat{z}$ sur $\hat{n}$, et Θ l'angle entre $\hat{r}$ et la direction définie par les angles (θ,ϕ) (figure 9.4)

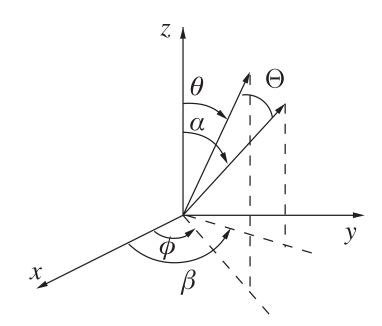


Fig. 9.4 – Configuration des angles dans (9.67).

$$\cos\Theta = \cos\alpha\cos\theta + \sin\alpha\sin\theta\cos(\beta-\phi)$$

L'angle Θ entre $\mathcal{R}^{-1}\hat{r}$ et l'axe des z est le même que l'angle entre $\hat{n}$ et $\hat{r}$. Il suffit alors de choisir $m = 0$ dans (9.65) pour obtenir

$$P_l(\cos\Theta) = \frac{4\pi}{2l+1}\sum_{m=-l}^{l}[Y_l^m(\theta,\phi)]^* Y_l^m(\alpha,\beta) \tag{9.67}$$

4. *Parité des harmoniques sphériques.* L'opérateur parité Π défini au § 7.3.3 agit sur une fonction d'onde $\psi(\vec{r}\,)$ suivant

$$[\Pi\psi](\vec{r}) = \psi(-\vec{r}) \tag{9.68}$$

Π commute avec le moment angulaire orbital $\vec{L}$, et plus généralement avec $\vec{J}$. En effet, la matrice représentative de l'opérateur parité dans l'espace à trois

dimensions $\mathbb{R}^3$ est la matrice $-I$, qui commute avec toute matrice de rotation $\mathcal{R}$, d'où

$$[U(\mathcal{R}),\Pi] = 0 \Longrightarrow [\vec{J},\Pi] = 0 \;\text{ et }\; [\vec{L},\Pi] = 0 \tag{9.69}$$

Ceci implique les équations

$$\begin{aligned}\vec{L}^{\,2}\Pi\, Y_l^m &= \Pi\vec{L}^{\,2}Y_l^m = l(l+1)\Pi\, Y_l^m \\ L_z\Pi\, Y_l^m &= \Pi L_z\, Y_l^m = m\Pi\, Y_l^m\end{aligned}$$

qui montrent que $\Pi\, Y_l^m$ est proportionnel à Y_l^m

$$\Pi Y_l^m = \alpha(l,m)Y_l^m$$

Y_l^m est donc fonction propre de Π et comme $\Pi^2 = I$, $\alpha(l,m) = \pm 1$. Montrons que $\alpha(l,m)$ est en fait indépendant de m en utilisant le fait que L_+ commute avec Π

$$\begin{aligned}L_+\Pi\, Y_l^m &= \alpha(l,m)L_+Y_l^m = \alpha(l,m)[l(l+1)-m(m+1)]^{1/2}\, Y_l^{m+1} \\ &= \Pi L_+Y_l^m = [l(l+1)-m(m+1)]^{1/2}\,\Pi\, Y_l^{m+1} \\ &= [l(l+1)-m(m+1)]^{1/2}\,\alpha(l,m+1)Y_l^{m+1}\end{aligned}$$

ce qui entraîne que $\alpha(l,m+1) = \alpha(l,m)$: $\alpha(l,m)$ est indépendant de m et

$$[\Pi\, Y_l^m](\hat{r}) = \alpha(l)Y_l^m(\hat{r}) = Y_l^m(-\hat{r})$$

La transformation $\hat{r} \to -\hat{r}$ correspond à

$$\theta \to \pi - \theta \qquad\qquad \phi \to \phi + \pi \tag{9.70}$$

Si $m = 0$, $Y_l^0 \propto P_l(\cos\theta)$; d'après (9.62) et compte tenu de $P_l(-u) = (-1)^l P_l(u)$, on trouve $\alpha(l) = (-1)^l$ et

$$Y_l^m(\theta,\phi) = (-1)^l Y_l^m(\pi-\theta,\phi+\pi) \;\text{ ou }\; Y_l^m(\hat{r}) = (-1)^l Y_l^m(-\hat{r}) \tag{9.71}$$

9.4 Particule dans un potentiel central

9.4.1 Équation d'onde radiale

Nous allons utiliser les résultats précédents pour montrer que l'équation de Schrödinger à trois dimensions, qui est une équation aux dérivées partielles, se ramène à une équation différentielle ordinaire lorsque le potentiel est central, c'est-à-dire invariant par rotation

$$V(\vec{r}\,) = V(|\vec{r}\,|) = V(r)$$

Dans ce cas, comme l'énergie cinétique est un opérateur scalaire, le hamiltonien total pour une particule de masse M

$$H = \frac{\vec{P}^{\,2}}{2M} + V(r) \tag{9.72}$$

est invariant par rotation : $[H, \vec{J}] = 0$. Dans notre problème, seul intervient le moment angulaire orbital, puisque les seuls opérateurs dont nous disposons sont $\vec{P}$ et $\vec{R}$

$$[H, \vec{L}] = 0 \quad \text{ou} \quad [H, L_x] = [H, L_y] = [H, L_z] = 0 \tag{9.73}$$

Dans l'espace $L^{(2)}_{\vec{r}}(\mathbb{R}^3)$, l'opérateur énergie cinétique est proportionnel au laplacien ∇^2

$$-\vec{P}^{\,2} = -(-\mathrm{i}\nabla)^2 = \nabla^2 = \frac{1}{r}\,\frac{\partial^2}{\partial r^2}\,r + \frac{1}{r^2}\,\left[\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\phi^2}\right] \tag{9.74}$$

en écrivant le laplacien en cordonnées sphériques. Comparant avec (9.49), on reconnaît dans la partie angulaire du laplacien l'opérateur $\vec{L}^{\,2}$

$$\nabla^2 = \frac{1}{r}\,\frac{\partial^2}{\partial r^2}r - \frac{1}{r^2}\,\vec{L}^{\,2} \tag{9.75}$$

Cette équation confirme la relation de commutation $[H, \vec{L}] = 0$, puisque $[\vec{L}^{\,2}, \vec{L}] = 0$ et que la partie radiale du laplacien, qui ne dépend pas des angles, commute manifestement avec $\vec{L}$. Nous pouvons donc écrire le hamiltonien (9.72)

$$H = -\frac{1}{2M}\,\frac{1}{r}\,\frac{\partial^2}{\partial r^2}\,r + \frac{1}{2Mr^2}\,\vec{L}^{\,2} + V(r) \tag{9.76}$$

Compte tenu de leurs relations de commutation, nous savons qu'il est possible de diagonaliser simultanément H, $\vec{L}^{\,2}$ et L_z. Soit $\psi_{lm}(\vec{r})$ une fonction propre commune à ces trois opérateurs. Comme il n'existe qu'un seul harmonique sphérique (l, m), si

$$\vec{L}^{\,2}\psi_{lm} = l(l+1)\psi_{lm} \quad \text{et} \quad L_z\psi_{lm} = m\psi_{lm}$$

alors ψ_{lm} doit être proportionnel[4] à Y_l^m

$$\psi_{lm}(r, \theta, \phi) = f_l(r)Y_l^m(\theta, \phi) = \frac{u_l(r)}{r}\,Y_l^m(\theta, \phi) \tag{9.77}$$

Il est commode de factoriser $1/r$: $u_l(r)$ est la *fonction d'onde radiale*. Examinons l'action de H sur ψ_{lm}

$$H\psi_{lm}(r, \theta, \phi) = \left[-\frac{1}{2M}\,\frac{1}{r}\,\frac{\partial^2}{\partial r^2}\,u_l(r) + \left(\frac{l(l+1)}{2Mr^2} + V(r)\right)\,\frac{u_l(r)}{r}\right]\,Y_l^m(\theta, \phi)$$

L'équation aux valeurs propres

$$H\psi_{lm} = E_l\psi_{lm}$$

4. Nous anticipons sur le fait, prouvé quelques lignes plus bas, que f_l ne dépend pas de m.

devient

$$\boxed{\left[-\frac{1}{2M}\frac{\mathrm{d}^2}{\mathrm{d}r^2}+\frac{l(l+1)}{2Mr^2}+V(r)\right]u_l(r)=E_l u_l(r)} \tag{9.78}$$

La fonction d'onde radiale et l'énergie sont affectées uniquement d'un indice l, et non m, car elles sont indépendantes de m d'après (9.78). *Chaque valeur de l'énergie aura donc −au moins− une dégénérescence* $(2l+1)$. Ceci aurait pu être prévu à partir de la relation de commutation $[H, L_\pm] = 0$: si

$$H\psi_{lm} = E_{lm}\psi_{lm},$$

en utilisant un raisonnement analogue à celui qui nous a permis de montrer que $\alpha(l, m)$ est indépendant de m, on déduit que E_{lm} est indépendant de m (exercice 9.7.7). Pour chaque valeur du moment angulaire l, ou pour chaque *onde partielle* l, nous avons réduit l'équation de Schrödinger à une équation différentielle ordinaire dans la seule variable r. Suivant une tradition historique, les ondes partielles sont étiquetées $s,\ p,\ d,\ f,\ g,\ h,\ \ldots$

$l = 0$: onde s ; $l = 1$: onde p ; $l = 2$: onde d ; $l = 3$: onde f

et ensuite par ordre alphabétique : $l = 4$: onde g, etc. Dans chaque onde partielle, l'équation (9.78) montre que le potentiel $V(r)$ doit être remplacé par un *potentiel effectif* $V_l(r)$ (figure 9.5)

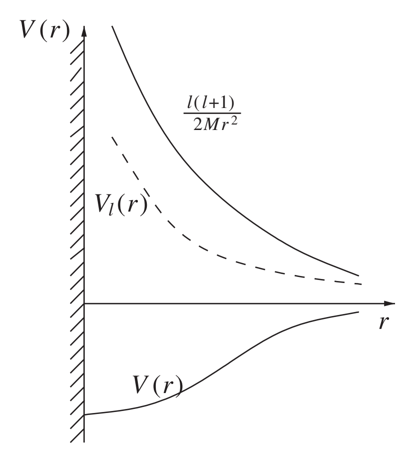


FIG. 9.5 – Potentiel effectif. Les lignes en traits pleins représentent les potentiels $V(r)$ et la barrière centrifuge $l(l+1)/(2mr^2)$ et les tirets leur somme, le potentiel effectif $V_l(r)$ dans l'onde partielle de moment angulaire l.

$$V_l(r) = V(r) + \frac{l(l+1)}{2Mr^2} \tag{9.79}$$

Le terme $l(l+1)/(2Mr^2)$ est appelé terme de *barrière centrifuge*. Ce terme est aussi présent en mécanique classique, où l'énergie peut s'écrire

$$E = \frac{1}{2}\, M\vec{v}^{\,2} + V(r) = \frac{1}{2}\, M(v_r^2 + \omega^2 r^2) + V(r)$$

v_r est la vitesse radiale et ω la vitesse angulaire. Comme[5] $l = M\omega r^2$ et que $\vec{l}$ est constant dans le cas d'une force centrale

$$E = \frac{1}{2}\, Mv_r^2 + \frac{l^2}{2Mr^2} + V(r) = \frac{1}{2}\, Mv_r^2 + V_l(r)$$

Le terme $l^2/(2Mr^2)$ correspond à la force centrifuge

$$-\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{l^2}{2Mr^2}\right) = \frac{l^2}{Mr^3} = M\omega^2 r$$

Ce terme tend à éloigner la particule du centre de force et correspond à un potentiel répulsif. En mécanique quantique, on remplacera pour chaque valeur de l l'opérateur $\vec{L}^{\,2}$ par sa valeur propre $l(l+1)$, et au potentiel $V(r)$ s'ajoutera un potentiel répulsif $l(l+1)/(2Mr^2)$.

Toutes les fonctions $\psi_{lm}(\vec{r})$ de type (9.77) avec $u_l(r)$ solution de (9.78) ne sont pas acceptables physiquement. En effet, si la fonction $\psi_{lm}(\vec{r})$ représente un état lié, elle doit obéir à la condition de normalisation

$$\int \mathrm{d}^3 r\, |\psi_{lm}(\vec{r})|^2 = 1 \qquad (9.80)$$

Si $\psi_{lm}(\vec{r})$ représente un état de diffusion, un comportement à l'infini en onde plane plus onde sphérique $\exp(\pm\mathrm{i}kr)/r$ est acceptable (*cf.* (9.81)). Dans le cas d'un état lié, l'équation (9.78) possède en général, à l fixé, plusieurs solutions : en effet, cette équation est identique à celle d'un problème unidimensionnel dans l'intervalle $]0, +\infty[$, car $0 < r < \infty$, le potentiel effectif étant le potentiel $V_l(r)$ (9.79). La fonction d'onde radiale et l'énergie sont étiquetées par un nombre quantique supplémentaire n' : $n' = 0,\ 1,\ 2,\ \ldots$ La fonction d'onde et l'énergie seront notées $u_{n'l}(r)$ et $E_{n'l}$. Si le potentiel $V(r)$ est suffisamment régulier, on montre que n' est égal au nombre de zéros, aussi appelé nombre de nœuds, de la fonction d'onde radiale $u_{n'l}(r)$ (*cf.* § 8.3.3). Le nombre quantique n' classe les valeurs de l'énergie par ordre croissant

$$n'_1 > n'_2 \ \Rightarrow\ E_{n'_1 l} > E_{n'_2 l}$$

Nous montrerons au chapitre 13 que les fonctions d'onde des états de diffusion sont étiquetées par le vecteur d'onde $\vec{k}$

$$r \to \infty \ :\ \psi_{\vec{k}}(\vec{r}) \simeq \mathrm{e}^{\,\mathrm{i}\vec{k}\cdot\vec{r}} + f(\theta, \phi)\frac{\mathrm{e}^{\,\pm\mathrm{i}kr}}{r} \qquad (9.81)$$

5. Suivant notre convention habituelle, les lettres minuscules désignent des quantités classiques (nombres), ou des nombres quantiques.

Il est possible d'analyser le comportement des fonctions d'onde $u_{n'l}(r)$ pour $r \to 0$. En effet, dans tous les cas d'intérêt physique, le terme de barrière centrifuge est le plus singulier lorsque $r \to 0$ et il contrôle le comportement de $u_l(r)$ dans cette limite. Si l'on suppose un comportement en loi de puissances[6]

$$r \to 0 \;:\; u_l(r) \propto r^\alpha$$

et que l'on reporte dans (9.78), on obtient pour les deux termes les plus singuliers en $r^{\alpha-2}$

$$-\frac{1}{2M}\,\alpha(\alpha-1)r^{\alpha-2} + \frac{l(l+1)}{2M}\,r^{\alpha-2} = 0$$

ce qui implique

$$\alpha(\alpha-1) = l(l+1)$$

soit $\alpha = l+1$ ou $\alpha = -l$. La seconde valeur est exclue car l'intégrale dans (9.80) divergerait à l'origine, sauf si $l = 0$. Cependant, pour $l = 0$, une solution $u_0(r) \propto$ cste, soit $\psi_l(\vec{r}) \propto 1/r$, bien que normalisable, n'est pas acceptable car elle ne peut pas être solution de l'équation de Schrödinger ; en effet

$$\nabla^2 \frac{1}{r} = -4\pi\delta(\vec{r})$$

En résumé, le comportement des fonctions d'onde radiales pour $r \to 0$ est

$$r \to 0 \;:\; u_l(r) \propto r^{l+1} \tag{9.82}$$

La fonction d'onde radiale s'annule à l'origine. Ceci peut se comprendre intuitivement : comme $0 < r < \infty$, tout se passe comme s'il y avait une barrière de potentiel infinie à $r = 0$, et nous savons que dans ce cas (voir § 8.3.2) la fonction d'onde doit s'annuler. Toutefois, les solutions en r^{-l} peuvent être utiles dans la résolution de l'équation de Schrödinger lorsque l'on se trouve dans une région où r est strictement positif.

L'exemple de l'atome d'hydrogène, étudié dans la sous-section suivante, conduit à redéfinir le nombre quantique radial (ou principal)

$$n' \to n = n' + l + 1 \tag{9.83}$$

9.4.2 Atome d'hydrogène

Les résultats de la sous-section précédente permettent d'aborder le calcul des niveaux d'énergie et des fonctions d'onde de l'atome d'hydrogène, un des rares problèmes physiques pour lesquels on dispose d'une solution analytique.

---

6. La loi de puissances donnant le comportement à l'origine ne dépend pas du nombre quantique n' ou de k, et nous supprimons la dépendance en n' ou en k.

La masse M dans (9.78) est la masse de l'électron m_e, ou plus exactement la masse réduite μ (exercice 7.5.6)

$$\mu = \frac{m_e m_p}{m_e + m_p} \simeq m_e \tag{9.84}$$

où m_p est la masse du proton. Nous conserverons néanmoins m_e dans les équations, afin de souligner l'ordre de grandeur (à 0.1 % près !) de la masse pertinente dans ce problème. Le potentiel $V(r)$ est le potentiel coulombien attractif entre l'électron et le proton

$$V(r) = -\frac{q_e^2}{4\pi\varepsilon_0 r} = -\frac{e^2}{r} \tag{9.85}$$

et l'équation (9.78) devient

$$\left[-\frac{1}{2m_e}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{l(l+1)}{2m_e r^2} - \frac{e^2}{r}\right] u_{nl}(r) = E_{nl}\, u_{nl}(r) \tag{9.86}$$

Il est toujours conseillé en physique de rendre les équations adimensionnelles par des changements de variables appropriés. Dans le présent problème, l'unité de longueur naturelle est le rayon de Bohr (1.41), $a_0 = 1/(m_e e^2)$ et l'unité d'énergie le Rydberg[7] (1.42), $R_\infty = e^2/(2a_0) = m_e e^4/2$. Ceci suggère de définir les quantités sans dimensions x et ε_{nl}

$$x = \frac{r}{a_0} = me^2 r \qquad \varepsilon_{nl} = -\frac{E_{nl}}{R_\infty} = -\frac{2a_0\, E_{nl}}{e^2} \tag{9.87}$$

Dans la suite du calcul, nous nous limitons aux états liés pour lesquels $E_{nl} < 0$ et donc $\varepsilon_{nl} > 0$, ce qui explique le choix du signe moins. Définissant également

$$v_{nl}(x) = u_{nl}(r) = u_{nl}(a_0 x)$$

nous obtenons après simplification par $(2m_e a_0^2)^{-1}$

$$\left[-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{l(l+1)}{x^2} - \frac{2}{x}\right] v_{nl}(x) = -\varepsilon_{nl}\, v_{nl}(x) \tag{9.88}$$

Nous nous contenterons de déterminer la solution dans le cas $l = 0$, c'est-à-dire dans l'onde s, en renvoyant le cas général à l'exercice 9.7.9. Afin d'alléger les notations, nous posons

$$v_{n0}(x) = v(x) \qquad \varepsilon_{n0} = \varepsilon$$

et (9.88) devient

$$\frac{\mathrm{d}^2 v(x)}{\mathrm{d}x^2} = \left(\varepsilon - \frac{2}{x}\right) v(x)$$

7. Rappelons que nous avons choisi un système d'unités où $\hbar = 1$: si l'on rétablit $\hbar$, $a_0 = \hbar^2/(m_e e^2)$ et $R_\infty = m_e e^4/(2\hbar^2)$.

D'après la sous-section précédente, nous savons que $v(x) \propto x$ pour $x \to 0$. Cherchons le comportement dominant pour $x \to \infty$ en négligeant le terme en $2/x$. Nous avons alors[8]

$$v(x) \sim \exp(\pm\sqrt{\varepsilon}\, x)$$

Le comportement en $\exp(\sqrt{\varepsilon}\, x)$ est inacceptable car la fonction d'onde ne serait pas normalisable, en raison de la divergence exponentielle. Le seul comportement possible est en $\exp(-\sqrt{\varepsilon}\, x)$. Afin de prendre en compte l'information contenue dans le comportement à l'infini, définissons une nouvelle fonction $f(x)$ par

$$v(x) = \mathrm{e}^{-\alpha x} f(x) \qquad \alpha^2 = \varepsilon$$

Ce changement de fonction transforme l'équation différentielle pour $v(x)$ en

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} - 2\alpha \frac{\mathrm{d}f}{\mathrm{d}x} + \frac{2}{x} f = 0 \tag{9.89}$$

Cherchons pour $f(x)$ un développement en puissances de x, sachant que $f(x) \propto x$ pour $x \to 0$

$$f(x) = \sum_{k=1}^{\infty} a_k x^k \tag{9.90}$$

L'équation (9.89) détermine une relation de récurrence sur les coefficients a_k

$$\sum_{k=1}^{\infty} k(k-1) a_k x^{k-2} - 2\alpha \sum_{k=1}^{\infty} k a_k x^{k-1} + 2 \sum_{k=1}^{\infty} a_k x^{k-1} = 0$$

En remarquant que pour $k = 1$ le premier terme de l'équation précédente s'annule et en réétiquetant k

$$\sum_{k=1}^{\infty} \left[k(k+1) a_{k+1} - 2(\alpha k - 1) a_k \right] x^{k-1} = 0 \tag{9.91}$$

L'annulation du coefficient de x^{k-1} donne une relation entre a_{k+1} et a_k

$$a_{k+1} = \frac{2(\alpha k - 1)}{k(k+1)}\, a_k$$

Si l'on fixe arbitrairement a_1, tous les a_k se déduisent de a_1. Pour $k \gg 1$, la relation de récurrence est approximativement

$$a_{k+1} \simeq \frac{2\alpha}{k}\, a_k \Rightarrow a_k \simeq \frac{(2\alpha)^k}{k!}\, a_1$$

et

$$\sum_{k=1}^{\infty} a_k x^k \sim \sum_{k=1}^{\infty} \frac{(2\alpha)^k}{k!}\, a_1 x^k \sim a_1\, \mathrm{e}^{2\alpha x}$$

8. En fait, ce comportement n'est déterminé qu'à un polynôme multiplicatif près.

Ce comportement entraîne pour $x \to \infty$

$$v(x) \sim \mathrm{e}^{2\alpha x}\mathrm{e}^{-\alpha x} \sim a_1 \mathrm{e}^{\alpha x}$$

qui rend la fonction d'onde non normalisable. La seule façon d'éviter la divergence exponentielle est que le développement (9.90) s'arrête pour une valeur entière $k = n$, ce qui ne peut se produire que si $\alpha n = 1$. Les valeurs possibles de ε sont donc étiquetées par un entier n

$$\varepsilon_n = \alpha^2 = \frac{1}{n^2}$$

et il en est de même pour celles de l'énergie

$$E_n = E_{n0} = -\frac{me^4}{2}\frac{1}{n^2} = -\frac{R_\infty}{n^2} \tag{9.92}$$

L'exercice 9.7.9 montre que les énergies possibles pour $l \neq 0$ sont de la forme

$$E_{nl} = -\frac{R_\infty}{n^2} \qquad n = \ l+1,\ \ l+2,\ \cdots \tag{9.93}$$

Les fonctions d'onde radiales $v_{n0}(x)$ pour $n = 1$ et 2 des états liés de l'atome d'hydrogène dans l'onde s, normalisées à un, sont données par

$$v_{10}(x) \ = \ 2x\,\mathrm{e}^{-x} \tag{9.94}$$

$$v_{20}(x) \ = \ \frac{1}{\sqrt{2}}\,x\left(1 - \frac{x}{2}\right)\mathrm{e}^{-x/2} \tag{9.95}$$

La fonction d'onde radiale de l'état $n = 2$, $l = 1$ (onde p) est

$$v_{21}(x) = \frac{1}{2\sqrt{6}}\ x^2\mathrm{e}^{-x/2} \tag{9.96}$$

En conclusion, nous avons obtenu le spectre de l'atome d'hydrogène (figure 9.6). On remarque que les niveaux sont dégénérés, sauf dans le cas $n = 1$. Pour une valeur donnée de n, toutes les valeurs de l comprises entre $l = 0$ et $l = n - 1$ sont possibles, et la dégénérescence est

$$G(n) = \sum_{l=0}^{n-1}(2l+1) = n^2$$

Cette dégénérescence est propre au potentiel coulombien. Le spectre de l'électron externe d'un alcalin est qualitativement semblable à celui de l'atome d'hydrogène (figure 9.7), mais il n'y a plus de dégénérescence. En mécanique classique, le potentiel coulombien présente aussi une particularité remarquable : il est le seul, avec le potentiel harmonique $V(r) \propto r^2$, pour lequel les trajectoires

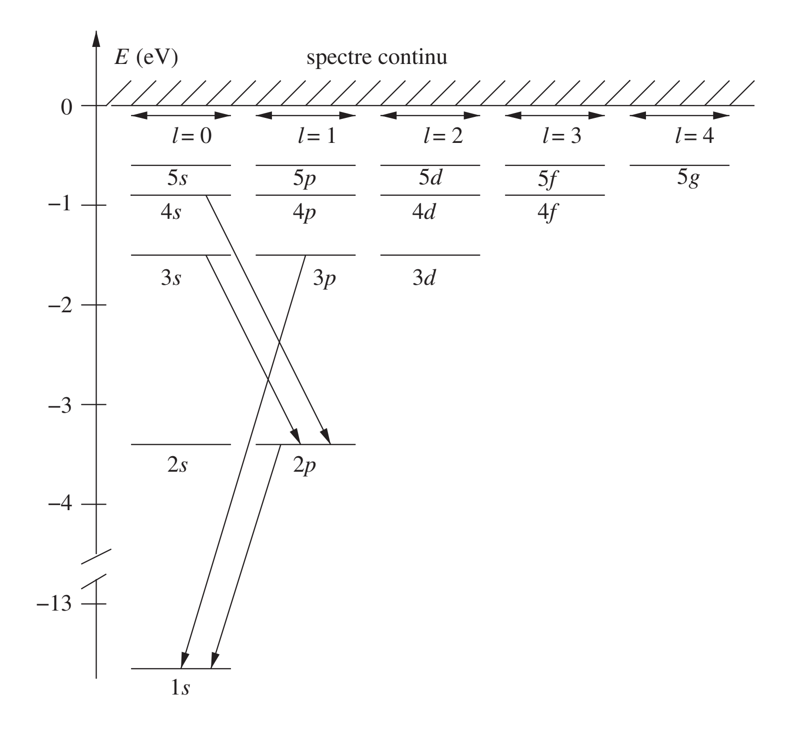


FIG. 9.6 – Spectre de l'atome d'hydrogène avec des exemples de transitions possibles.

se referment sur elles-mêmes[9]. Dans les deux cas, cette propriété du mouvement classique, ainsi que les dégénérescences associées du problème quantique, ont pour origine une symétrie supplémentaire. Cette symétrie conduit à une loi de conservation supplémentaire, celle du vecteur de Lenz dans le cas coulombien.

9.5 Distributions angulaires des désintégrations

9.5.1 Rotations de π, parité, réflexion par rapport à un plan

Dans cette section, nous nous proposons d'étudier les désintégrations d'une particule C en deux particules A et B

$$C \to A + B \tag{9.97}$$

9. Les deux comportements peuvent être reliés : *cf.* Basdevant et Dalibard [2001], chapitre 11, exercice 3.

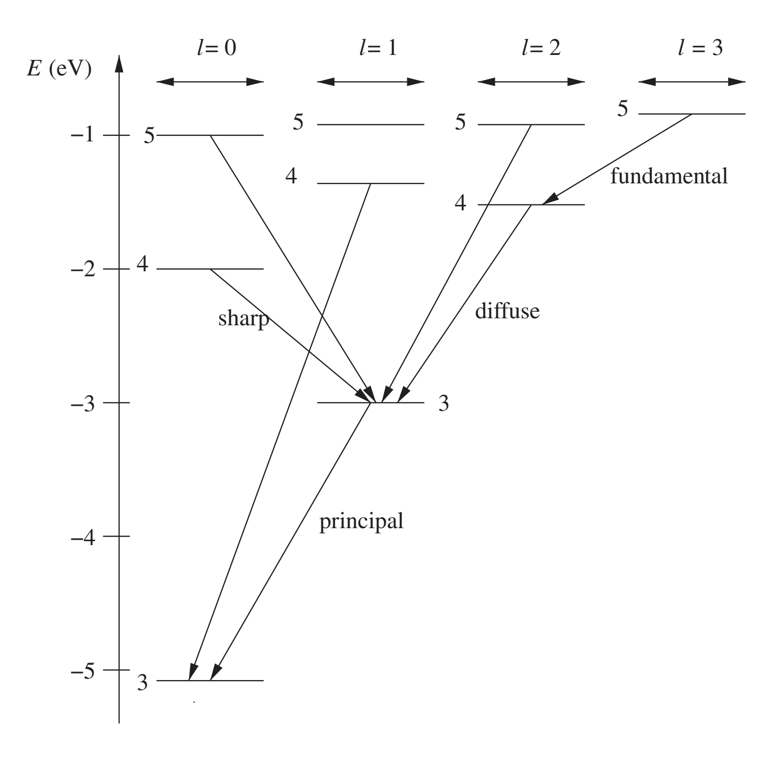


FIG. 9.7 – Spectre de l'atome de sodium avec des exemples de transitions possibles.

Nous choisissons un référentiel d'inertie où la particule C est au repos ; les particules A et B partent donc avec des impulsions opposées $\vec{p}$ et $-\vec{p}$, respectivement. Les processus (9.97) comprennent les *désintégrations (ou transitions) radiatives*, avec émission d'un photon : un niveau excité A^* d'un atome, d'une molécule ou d'un noyau atomique émet un photon γ tandis que le système passe dans un niveau d'énergie inférieure A, qui peut être ou non l'état fondamental

$$A^* \to A + \gamma \tag{9.98}$$

Les états A^* et A peuvent aussi correspondre à des particules différentes, comme par exemple dans la désintégration (exercice 9.7.17)

$$\Sigma^0 \to \Lambda^0 + \gamma \tag{9.99}$$

où les particules Σ^0 et Λ^0 sont des particules neutres formées d'un quark up, un quark down et un quark étrange.

L'invariance par rotation du hamiltonien responsable de la désintégration entraîne la conservation du moment angulaire, ce qui conduit à des contraintes sur les amplitudes de désintégration et à des conséquences importantes sur la distribution angulaire des particules finales. Si le hamiltonien responsable de

la désintégration est invariant par parité, ce qui est le cas pour les interactions électromagnétiques et fortes, mais non pour les interactions faibles, on obtiendra des contraintes supplémentaires. Il est commode d'introduire l'opérateur $\mathcal{Y}$, produit d'une rotation de π autour de l'axe Oy et de l'opération parité Π (§ 7.3.3)

$$Y = \mathrm{e}^{-\mathrm{i}\pi J_y} \qquad \mathcal{Y} = Y\,\Pi = \mathrm{e}^{-\mathrm{i}\pi J_y}\,\Pi = \Pi\mathrm{e}^{-\mathrm{i}\pi J_y} \tag{9.100}$$

Cette opération n'est autre qu'une réflexion par rapport au plan xOz : $\mathcal{Y}$ est l'opérateur de réflexion par rapport à ce plan. Étudions tout d'abord l'action de Y. Cet opérateur transforme J_x en $-J_x$, J_z en $-J_z$ et laisse J_y inchangé

$$Y^{-1}J_zY = -J_z \qquad Y^{-1}J_\pm Y = -J_\mp \tag{9.101}$$

Examinons l'action de Y sur un état $|jm\rangle$

$$J_z(Y|jm\rangle) = -YJ_z|jm\rangle = -m(Y|jm\rangle)$$

L'état $Y|jm\rangle$ est donc égal à $|j,-m\rangle$ à un facteur de phase près

$$Y|jm\rangle = \mathrm{e}^{\mathrm{i}\alpha(j,m)}\,|j,-m\rangle$$

car Y est unitaire et conserve la norme. Ce résultat n'est pas surprenant car l'action de Y équivaut à renverser la direction de l'axe de quantification du moment angulaire. Suivant une stratégie déjà mise en œuvre dans le cas de la parité, on utilise l'action de J_+ pour relier $\alpha(j,m)$ à $\alpha(j,m+1)$

$$\begin{aligned}
J_+Y|jm\rangle &= \mathrm{e}^{\mathrm{i}\alpha(j,m)}J_+|j,-m\rangle = \sqrt{j(j+1)-m(m-1)}\,\mathrm{e}^{\mathrm{i}\alpha(j,m)}\;|j,-m+1\rangle \\
&= -YJ_-|jm\rangle = -\sqrt{j(j+1)-m(m-1)}\;Y|j,m-1\rangle \\
&= -\sqrt{j(j+1)-m(m-1)}\;\mathrm{e}^{\mathrm{i}\alpha(j,m-1)}\,|j,-m+1\rangle
\end{aligned}$$

soit

$$\mathrm{e}^{\mathrm{i}\alpha(j,m-1)} = -\mathrm{e}^{\mathrm{i}\alpha(j,m)}$$

Comme Y est une rotation de π, Y^2 est une rotation de 2π, $Y^2 = (-1)^{2j}$ et

$$Y^2|jm\rangle = \mathrm{e}^{\mathrm{i}\alpha(j,m)}\mathrm{e}^{\mathrm{i}\alpha(j,-m)}|jm\rangle = \mathrm{e}^{2\mathrm{i}\alpha(j,m)}(-1)^{2m}|jm\rangle = (-1)^{2j}|jm\rangle$$

d'où les deux solutions possibles

$$\mathrm{e}^{\mathrm{i}\alpha(j,m)} = (-1)^{j-m} \;\text{ ou }\; \mathrm{e}^{\mathrm{i}\alpha(j,m)} = (-1)^{j+m}$$

Les deux solutions sont identiques pour j entier, et pour $j = 1/2$ on vérifie sur (9.38) que la première solution est la bonne ; on peut montrer que c'est aussi le cas pour tout j demi-entier. En fin de compte,

$$Y|jm\rangle = (-1)^{j-m}|j,-m\rangle \qquad Y^{-1}|jm\rangle = (-1)^{j+m}|j,-m\rangle \tag{9.102}$$

9.5.2 Transitions dipolaires

Nous examinons dans cette sous-section les transitions radiatives du type (9.98). Revenons tout d'abord sur la description de la polarisation d'un photon que nous avions étudiée au chapitre 3, en la replaçant dans le cadre général du moment angulaire. Nous avions déterminé le générateur infinitésimal des rotations de la polarisation lorsque la rotation se faisait autour de la direction de propagation, prise comme axe Oz. Dans la base des états de polarisation linéaire $|x\rangle$ et $|y\rangle$, ce générateur infinitésimal est donné par (3.26)

$$\Sigma_z = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}$$

En effet, nous avons vu en (3.29) que $\exp(-\mathrm{i}\theta\Sigma_z)$ effectue une rotation d'angle θ de la polarisation dans le plan xOy, et on peut identifier Σ_z et la composante z du moment angulaire du photon : $\Sigma_z = J_z$. L'action de l'opérateur $\exp(-\mathrm{i}\theta\Sigma_z)$ sur les états de polarisation circulaire droite $|D\rangle$ et gauche $|G\rangle$ (3.11) est d'après (3.29)

$$\exp(-\mathrm{i}\theta\Sigma_z)|D\rangle = \mathrm{e}^{-\mathrm{i}\,\theta}|D\rangle \qquad \exp(-\mathrm{i}\theta\Sigma_z)|G\rangle = \mathrm{e}^{\mathrm{i}\,\theta}|G\rangle$$

ce qui prouve que les états $|D\rangle$ et $|G\rangle$ ont des nombres quantiques magnétiques $m = 1$ et $m = -1$, respectivement[10]. Par ailleurs, la description du champ électromagnétique par un potentiel vecteur montre que le photon a un caractère vectoriel, et donc un spin 1, ce qui permet l'identification de $|D\rangle$ et $|G\rangle$ avec des états $|jm\rangle$ (figure 9.8)

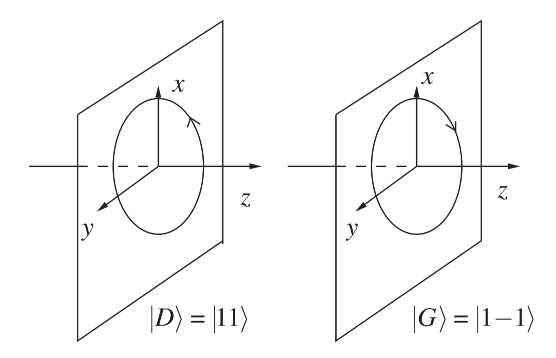


FIG. 9.8 – (a) Polarisation circulaire droite ; (b) polarisation circulaire gauche.

$$|D\rangle = |j = 1, m = 1\rangle = |11\rangle \qquad |G\rangle = |j = 1, m = -1\rangle = |1, -1\rangle \quad (9.103)$$

l'axe de quantification Oz du moment angulaire étant pris le long de la direction de propagation du photon. La valeur de m est appelé *l'hélicité* du

10. Un argument équivalent consiste à remarquer que $\Sigma_z|D\rangle = |D\rangle$ et que $\Sigma_z|G\rangle = -|G\rangle$.

photon : $m = +1$ correspond à une hélicité positive, $m = -1$ à une hélicité négative. Comme à un moment angulaire 1 correspondent trois valeurs possibles du nombre quantique magnétique, $m = +1$, $m = 0$, $m = -1$, on peut se demander ce que devient la valeur $m = 0$ dans le cas du photon. Une analyse générale due à Wigner (section 19.2) montre que pour une particule de masse nulle et de spin j, les seules valeurs propres permises de J_z sont $m = j$ et $m = -j$, l'axe Oz étant pris le long de la direction de propagation de la particule. Lorsque la parité n'est pas une symétrie du hamiltonien, les deux valeurs possibles sont indépendantes : si le neutrino avait une masse nulle[11], le neutrino, particule de spin $j = 1/2$, aurait toujours $m = -1/2$, et l'antineutrino, particule *différente*, $m = +1/2$. Les interactions du photon, qui sont les interactions électromagnétiques, conservent la parité, ce qui fait que la même particule peut avoir $m = 1$ et $m = -1$.

Reste à vérifier que la définition (9.103) correspond bien à une base standard du moment angulaire. Nous allons utiliser l'opérateur $Y = \exp(-\mathrm{i}\pi J_y)$ qui change la direction de propagation du photon, l'axe Oz restant inchangé. Son action sur les états de polarisation linéaire est (figure 9.9)

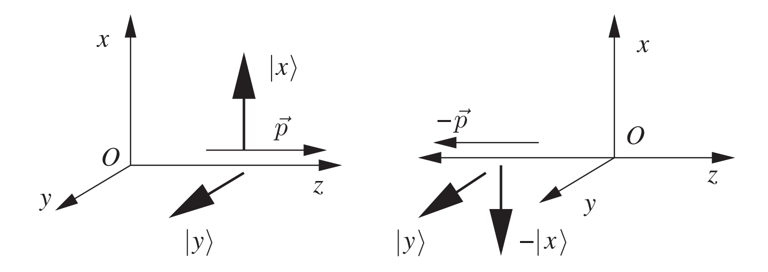


Fig. 9.9 – Action de Y sur les états de polarisation linéaire.

$$Y|x\rangle = -|x\rangle \qquad Y|y\rangle = |y\rangle$$

On en déduit l'action sur les états de polarisation circulaire $|D\rangle$ et $|G\rangle$ (3.11)

$$Y|D\rangle = Y\left[\frac{-1}{\sqrt{2}}\Big(|x\rangle + \mathrm{i}|y\rangle\Big)\right] = \frac{1}{\sqrt{2}}\Big(|x\rangle - \mathrm{i}|y\rangle\Big) = |G\rangle \tag{9.104}$$

La phase relative des états $|D\rangle$ et $|G\rangle$ correspond bien à celle d'une base standard puisque d'après (9.102)

$$Y|D\rangle = Y|1,1\rangle = (-1)^{1-1}|1,-1\rangle = |G\rangle$$

Le choix (3.11) est aussi confirmé par le fait que $|D\rangle$ et $|G\rangle$ sont donnés par les mêmes combinaisons que les composantes sphériques $\hat{r}_1$, $\hat{r}_{-1}$ et $\hat{r}_0$ (9.64) de $\hat{r}$.

11. Ce qui a longtemps semblé probable, mais a été contredit par l'expérience : voir l'exercice 4.4.6. et la note 4 du chapitre 1.

Revenant à (9.98), appelons $\vec{p}$ l'impulsion du photon, que nous choisissons comme axe Oz, soit $|jm\rangle$ l'état de moment angulaire de A^* – on dit souvent que l'état excité a un spin j –, $|j'm'\rangle$ l'état de moment angulaire du niveau final A, ou spin du niveau final A, et $|1\mu\rangle$ celui du photon. En raison de l'invariance par rotation, le moment angulaire est conservé dans la transition

$$\vec{J} = \vec{J'} + \vec{S} + \vec{L}$$

où $\vec{S}$ est le spin du photon et $\vec{L}$ le moment angulaire orbital. La projection de cette équation sur Oz donne

$$m = m' + \mu + m_l$$

Il est facile de se convaincre que le nombre quantique magnétique du moment angulaire orbital est nul : $m_l = 0$. En effet, la fonction d'onde spatiale du photon est une onde plane

$$\mathrm{e}^{\,\mathrm{i}\vec{p}\cdot\vec{r}} = \mathrm{e}^{\,\mathrm{i}pz}$$

qui est invariante par rotation autour de Oz : la composante suivant Oz du moment angulaire orbital doit être nulle. On peut aussi remarquer que d'après (9.47)

$$L_z \mathrm{e}^{\,\mathrm{i}pz} = -\mathrm{i}\frac{\partial}{\partial\phi}\,\mathrm{e}^{\,\mathrm{i}pr\cos\theta} = 0$$

La conservation du moment angulaire suivant Oz donne donc

$$\begin{array}{rl} \text{photon droit final} & m = m' + 1 \\ \text{photon gauche final} & m = m' - 1 \end{array} \qquad (9.105)$$

Si A et A^* ont un spin nul ($j = j' = 0$), alors $m = m' = 0$ et les équations (9.105) n'ont pas de solution : il n'y a pas de transition radiative à un seul photon $j = 0 \to j' = 0$, souvent appelée transition $0 \to 0$. Les transitions radiatives $0 \to 0$ ne sont possibles qu'avec l'émission d'au moins deux photons, et la probabilité d'une telle transition est défavorisée par une puissance de la constante de structure fine $\alpha \simeq 1/137$.

Un cas plus intéressant, et qui se présente souvent en pratique, est celui où $j = 1$ et $j' = 0$. Si le photon est émis suivant Oz avec une hélicité $\mu = \pm 1$, deux cas sont possibles, compte tenu de $j' = m' = 0$

$$\text{photon droit final} \qquad m = 1 \qquad \mu = 1 \qquad (9.106)$$

$$\text{photon gauche final} \qquad m = -1 \qquad \mu = -1 \qquad (9.107)$$

Soit a l'amplitude de probabilité de (9.106), b celle de (9.107). Il faut bien comprendre qu'il s'agit de l'amplitude de probabilité *d'une transition*, analogue à celle calculée en (8.149), et non des amplitudes de probabilité telles qu'elles sont définies dans le postulat **II** du chapitre 4. Le module carré d'une amplitude de transition donne la probabilité de transition par unité de temps.

Les amplitudes a et b peuvent être vues comme les éléments de matrice d'un opérateur T, appelé *matrice de transition* (§ 13.4.2) calculable, au moins formellement, en fonction du hamiltonien et qui a les mêmes symétries que le hamiltonien. Anticipant sur ce qui va suivre, on définit un angle θ qui est l'angle entre la direction d'émission prise dans le plan xOz du photon et l'axe Oz, et on écrit les amplitudes de transition a et b (dans (9.105) $m' = 0$ puisque $j' = 0$))

$$\begin{aligned} a &= \langle D, \theta = 0|T|j = 1, m = 1\rangle = \langle D, \theta = 0|T|11\rangle \\ b &= \langle G, \theta = 0|T|j = 1, m = -1\rangle = \langle G, \theta = 0|T|1, -1\rangle \end{aligned} \tag{9.108}$$

Si la parité est une symétrie du hamiltonien responsable de la transition, T commute avec $\mathcal{Y}$ (9.100) : comme les deux amplitudes a et b correspondent à des transitions qui se déduisent l'une de l'autre par une réflexion par rapport au plan xOz (figure 9.10a et b), on doit avoir $|a| = |b|$. Pour déterminer la phase dans cette relation, on utilise

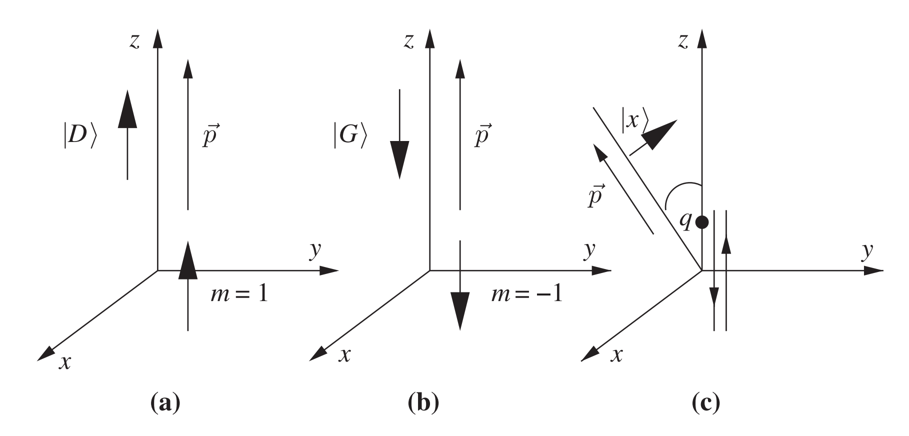


Fig. 9.10 – Émission de photons avec $\vec{p} \parallel Oz$. Les amplitudes (a) et (b) se déduisent l'une de l'autre par réflexion par rapport au plan xOz. (c) Polarisation linéaire du photon final. La charge q effectue un mouvement oscillant le long de Oz.

$$\begin{aligned} a &= \langle D, \theta = 0|\mathcal{Y}^{-1}T\mathcal{Y}|1, 1\rangle = \eta_\gamma\eta_A\eta_{A^*}\langle G, \theta = 0|T|1, -1\rangle \\ &= \eta_\gamma\eta_A\eta_{A^*}\, b \end{aligned} \tag{9.109}$$

où $\eta_X = \pm 1$ est la parité d'une particule X : si la particule X a une impulsion $\vec{p}$ et que l'on écrit son vecteur d'état $|X, \vec{p}\rangle$

$$\Pi|X, \vec{p}\rangle = \eta_X|X, -\vec{p}\rangle \tag{9.110}$$

La description du champ électromagnétique par un potentiel vecteur, qui est un vecteur polaire, montre que la parité du photon est $\eta_\gamma = -1$. Soit $\eta = \eta_A\eta_{A^*}$; il y a deux cas possibles

1. $\eta = -1 \qquad a = b$
2. $\eta = +1 \qquad a = -b$

Nous allons montrer que le premier cas est celui d'une transition dipolaire électrique, le second celui d'une transition dipolaire magnétique[12], par comparaison avec le cas classique le plus simple, le rayonnement d'une charge se déplaçant le long de Oz avec un mouvement harmonique. Le moment angulaire classique de cette charge par rapport à l'origine, et en particulier sa composante suivant z, est toujours nul, et le cas quantique se rapprochant le plus de cette situation est celui où l'état excité A^* possède un moment angulaire nul suivant Oz, c'est-à-dire qu'il est dans l'état $|j = 1, m = 0\rangle$. Afin de comparer la distribution angulaire du photon avec celle du rayonnement classique, il nous faut envisager le cas où l'angle d'émission du photon $\theta \neq 0$, l'état initial de l'atome étant $|10\rangle$. On obtient l'état $|D, \theta\rangle$ (resp. $|G, \theta\rangle$) du photon par une rotation d'angle θ autour de Oy à partir de $|D, \theta = 0\rangle$ (resp. $|G, \theta = 0\rangle$)

$$\begin{aligned} |D, \theta\rangle &= U[\mathcal{R}_{\hat{y}}(\theta)]|D, \theta = 0\rangle \\ |G, \theta\rangle &= U[\mathcal{R}_{\hat{y}}(\theta)]|G, \theta = 0\rangle \end{aligned}$$

L'amplitude d'émission dans la direction θ, par exemple pour un photon droit et un état initial $|j = 1, m = 0\rangle$, est

$$\begin{aligned} a_D^{m=0}(\theta) &= \langle D, \theta\,|T|10\rangle = \langle D, \theta = 0|U^\dagger[\mathcal{R}_{\hat{y}}(\theta)]T|10\rangle \\ &= \langle D, \theta = 0|TU^\dagger[\mathcal{R}_{\hat{y}}(\theta)]|10\rangle \\ &= \langle D, \theta = 0|T|11\rangle\langle 11|U^\dagger[\mathcal{R}_{\hat{y}}(\theta)]|10\rangle \\ &= a\, d_{01}^{(1)}(\theta) = \frac{a}{\sqrt{2}}\sin\theta \end{aligned} \tag{9.111}$$

Nous avons utilisé l'invariance par rotation de T, introduit un ensemble complet d'états intermédiaires dans le sous-espace $j = 1$, $\sum_m |1m\rangle\langle 1m|$ et obtenu l'élément de matrice de rotation grâce à (9.39). Un calcul analogue donne pour l'émission d'un photon gauche

$$a_G^{m=0}(\theta) = b\, d_{0-1}^{(1)}(\theta) = -\frac{b}{\sqrt{2}}\sin\theta \tag{9.112}$$

Si la polarisation finale est linéaire, on peut la décomposer sur les états[13] $|x\rangle$ polarisés dans le plan xOz et $|y\rangle$ polarisés suivant Oy (figure 9.10c)

$$|x\rangle = \frac{1}{\sqrt{2}}\Big(-|D\rangle + |G\rangle\Big) \qquad\qquad |y\rangle = \frac{\mathrm{i}}{\sqrt{2}}\Big(|D\rangle + |G\rangle\Big) \tag{9.113}$$

12. Ce résultat dépend des conventions de signe utilisées pour les états $|D\rangle$ et $|G\rangle$; on trouve le signe opposé dans Feynman *et al.* [1965], volume III, section 18.1, en raison d'une convention de signe différente dans la définition de $|D\rangle$.

13. Les états $|x\rangle$ et $|y\rangle$ sont définis par rapport à la direction $\vec{p}$ de propagation : figure 9.10 c.

et on déduit

$$\begin{aligned} a_x^{m=0}(\theta) &= \langle x,\theta|T|10\rangle = -\frac{1}{2}(a+b)\sin\theta \\ a_y^{m=0}(\theta) &= \langle y,\theta|T|10\rangle = -\frac{\mathrm{i}}{2}(a-b)\sin\theta \end{aligned} \qquad (9.114)$$

Dans le cas dipolaire électrique $a = b$, les photons sont polarisés suivant Ox ; dans le cas dipolaire magnétique, ils sont polarisés suivant Oy. C'est bien ce qui correspond au cas classique : si l'on prend, par exemple, une charge animée d'un mouvement d'oscillation harmonique suivant Oz avec un moment angulaire nul suivant Oz, le rayonnement est polarisé dans le plan xOz. Au contraire, un dipôle magnétique donnerait un rayonnement polarisé suivant Oy. Une transition dipolaire électrique correspond à $\eta = -1$, et donc *à un état initial et final ayant des parités opposées*, tandis qu'une transition dipolaire magnétique correspond à un état initial et un état final ayant la même parité. Dans les deux cas, la distribution angulaire est en $\sin^2\theta$.

9.5.3 Désintégrations : cas général

Revenons à la désintégration générale à deux corps (9.97), en appelant j_A, j_B et j_C les spins des particules A, B et C. Définissons les amplitudes de transition pour un état initial $|j_C m_C\rangle$ de la particule C vers les états finaux $|j_A m_A\rangle$ et $|j_B m_B\rangle$ des particules A et B, et supposons la particule A émise avec une impulsion $\vec{p}$ dirigée suivant une direction (θ, ϕ)

$$a^{m_C}_{m_A m_B}(\theta,\phi) = \langle m_A m_B; (\theta,\phi)|T|m_C\rangle \qquad (9.115)$$

Si la particule A est émise dans la direction $\hat{p} = (\theta, \phi)$, l'état

$$|m_A m_B; (\theta,\phi)\rangle = U(\mathcal{R})|m_A m_B; (\theta = 0, \phi = 0)\rangle$$

est le transformé de $|m_A m_B; (\theta = 0, \phi = 0)\rangle$ par la rotation (θ, ϕ) qui amène l'axe Oz sur la direction $\hat{p}$. Il faut souligner que, dans cet état, nous avons choisi l'axe de quantification du moment angulaire suivant $\hat{p}$, et que m_A et m_B sont les valeurs propres de $(\vec{J} \cdot \hat{p})$, et non J_z (figure 9.11). Lorsque la particule A est émise dans la direction Oz : $\theta = \phi = 0$, la conservation de la composante z du moment angulaire implique, comme dans le cas de la sous-section précédente, que $m_C = m_A + m_B$. Les seules amplitudes de transition non nulles sont

$$b_{m_A,m_B} = \langle m_A m_B; (\theta = 0, \phi = 0)|T|m_C = m_A + m_B\rangle \qquad (9.116)$$

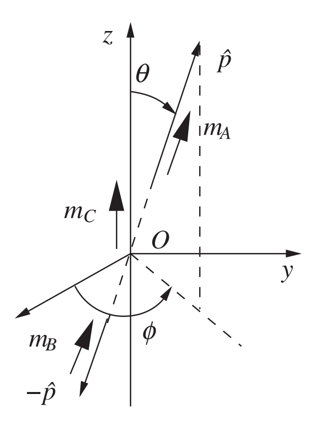


FIG. 9.11 – Désintégration $C \to A + B$.

En reprenant le raisonnement utilisé dans (9.111)

$$\begin{aligned}
a^{m_C}_{m_A,m_B}(\theta,\phi) &= \langle m_A m_B; (\theta,\phi)|T|m_C\rangle \\
&= \langle m_A, m_B; (\theta=0,\phi=0)|U^\dagger(\mathcal{R})\,T|m_C\rangle \\
&= \langle m_A, m_B; (\theta=0,\phi=0)|T|m'_C = m_A+m_B\rangle \\
&\quad \times\langle m'_C = m_A+m_B|U^\dagger(\mathcal{R})|m_C\rangle \\
&= b_{m_A,m_B}\left(D^{(j_C)}_{m_C;m_A+m_B}(\theta,\phi)\right)^* \qquad (9.117)\\
&= b_{m_A,m_B}\, d^{(j_C)}_{m_C;m_A+m_B}(\theta)\,\mathrm{e}^{\,\mathrm{i}m_C\phi} \qquad (9.118)
\end{aligned}$$

Si la parité est conservée dans la désintégration

$$\begin{aligned}
b_{m_A,m_B} &= \langle m_A, m_B; (\theta=0,\phi=0)|\mathcal{Y}^\dagger T\mathcal{Y}|m_C = m_A+m_B\rangle \\
&= \eta\,(-1)^{j_C-j_A-j_B}\, b_{-m_A,-m_B} \qquad (9.119)
\end{aligned}$$

où $\eta = \eta_A\eta_B\eta_C$ est le produit des parités des trois particules. La conservation de la parité divise par deux le nombre d'amplitudes indépendantes. Les amplitudes définies dans (9.118) sont appelées *amplitudes d'hélicité*. Toutefois, il faut prendre garde au fait que l'axe de quantification du moment angulaire de la particule B est souvent pris suivant la direction de son impulsion, soit $-\hat{p}$, ce qui fait que $m_B \to -m_B$. Les nombres quantiques magnétiques m_A et $-m_B$ (avec notre définition) sont les hélicités des particules A et B.

9.6 Composition de deux moments angulaires

9.6.1 Composition de deux spins 1/2

Nous avons construit au § 2.4.2 l'espace à quatre dimensions $\mathcal{H}_1 \otimes \mathcal{H}_2$ produit tensoriel de deux espaces à deux dimensions ; nous allons utiliser

cette construction en prenant comme espaces à deux dimensions ceux de deux spins 1/2, $\vec{S}_1$ et $\vec{S}_2$. Une base possible de $\mathcal{H}_1 \otimes \mathcal{H}_2$ est formée des vecteurs propres $|\varepsilon_1\varepsilon_2\rangle$, $\varepsilon = \pm$, de S_{1z} et de S_{2z}

$$|++\rangle\ ,\quad |+-\rangle\ ,\quad |-+\rangle\ \text{ et }\ |--\rangle \tag{9.120}$$

Les propriétés physiques diagonales dans cette base sont $\vec{S}_1^2$, $\vec{S}_2^2$, S_{1z} et S_{2z}

$$\vec{S}_1^2|\varepsilon_1\varepsilon_2\rangle = \frac{3}{4}\,|\varepsilon_1\varepsilon_2\rangle \qquad S_{1z}|\varepsilon_1\varepsilon_2\rangle = \varepsilon_1|\varepsilon_1\varepsilon_2\rangle \tag{9.121}$$

$$\vec{S}_2^2|\varepsilon_1\varepsilon_2\rangle = \frac{3}{4}\,|\varepsilon_1\varepsilon_2\rangle \qquad S_{2z}|\varepsilon_1\varepsilon_2\rangle = \varepsilon_2|\varepsilon_1\varepsilon_2\rangle \tag{9.122}$$

Cette base correspond au choix suivant d'un ensemble complet d'opérateurs compatibles : $\{\vec{S}_1^2, \vec{S}_2^2, S_{1z}, S_{2z}\}$. Il est possible de former une autre base intéressante en considérant le moment angulaire total $\vec{S}$ obtenu en additionnant $\vec{S}_1$ et $\vec{S}_2$

$$\vec{S} = \vec{S}_1 + \vec{S}_2 \tag{9.123}$$

$\vec{S}$ est bien le moment angulaire total, car il permet de construire le générateur infinitésimal dans l'espace produit tensoriel $\mathcal{H}_1 \otimes \mathcal{H}_2$ d'une rotation $\mathcal{R}_{\hat{n}}(\theta)$ d'angle θ autour d'un axe $\hat{n}$

$$U(\mathcal{R}_{\hat{n}}(\theta)) = \mathrm{e}^{-\mathrm{i}\theta(\vec{S}_1\cdot\hat{n})}\,\mathrm{e}^{-\mathrm{i}\theta(\vec{S}_2\cdot\hat{n})} = \mathrm{e}^{-\mathrm{i}\theta(\vec{S}\cdot\hat{n})} \tag{9.124}$$

où nous avons utilisé $[\vec{S}_1, \vec{S}_2] = 0$. Comme $\vec{S}_1^2$ et $\vec{S}_2^2$ sont des opérateurs scalaires, ils commutent avec $\vec{S}$, et un autre ensemble d'opérateurs compatibles est $\{\vec{S}_1^2, \vec{S}_2^2, \vec{S}^{\,2}, S_z\}$; nous vérifierons ultérieurement que cet ensemble est aussi complet. Cherchons les vecteurs de base de ce nouvel ensemble. Posant $|1,1\rangle = |++\rangle$, on vérifie

$$\begin{aligned} S_z|1,1\rangle &= |1,1\rangle \\ S_+|1,1\rangle &= (S_{1+} + S_{2+})|++\rangle = 0 \\ S_-|1,1\rangle &= (S_{1-} + S_{2-})|++\rangle = |+-\rangle + |-+\rangle = \sqrt{2}\ |1,0\rangle \end{aligned}$$

Cette dernière équation définit le vecteur unitaire $|1,0\rangle$, qui vérifie

$$S_z|1,0\rangle = (S_{1z} + S_{2z})\frac{1}{\sqrt{2}}\Big(|+-\rangle + |-+\rangle\Big) = 0$$

Enfin

$$\begin{aligned} S_-|1,0\rangle &= (S_{1-} + S_{2-})\frac{1}{\sqrt{2}}\Big(|+-\rangle + |-+\rangle\Big) = \sqrt{2}\ |--\rangle = \sqrt{2}\ |1,-1\rangle \\ S_z|1,-1\rangle &= -|1,-1\rangle \qquad\qquad S_-|1,-1\rangle = 0 \end{aligned}$$

Ces équations montrent que les trois vecteurs

$$\{|1,1\rangle,\ |1,0\rangle,\ |1,-1\rangle\}$$

forment une base standard pour un moment angulaire 1. Il suffit de vérifier les propriétés de la base standard pour S_z et S_-, car $S_+ = S_-^\dagger$ et $\vec{S}^{\,2} = \frac{1}{2}(S_+S_- + S_-S_+) + S_z^2$. Le calcul ci-dessus montre que l'on a bien construit une base standard, par exemple

$$S_-|1,1\rangle = \sqrt{j(j+1) - m(m-1)}\;|1,0\rangle = \sqrt{2}\;|1,0\rangle$$

Il faut enfin, pour obtenir une base de $\mathcal{H}_1 \otimes \mathcal{H}_2$, construire un quatrième vecteur orthogonal aux trois autres

$$|0,0\rangle = \frac{1}{\sqrt{2}}\Big(|+-\rangle - |-+\rangle\Big)$$

Ce vecteur n'est autre que le vecteur $|\Phi\rangle$ (2.63). Ce vecteur étant invariant par rotation correspond à un moment angulaire zéro, et on peut vérifier explicitement que

$$S_z|0,0\rangle = 0 \qquad S_\pm|0,0\rangle = 0$$

En résumé, *en composant deux moments angulaires 1/2, nous avons obtenu un moment angulaire* $s = 1$ *et un moment angulaire* $s = 0$. Une base standard de $\vec{S}^{\,2}$ et de S_z est formée des vecteurs correspondant à $s = 1$

$$s = 1 \qquad \begin{cases} |1,1\rangle = |++\rangle \\ |1,0\rangle = \frac{1}{\sqrt{2}}\Big(|+-\rangle + |-+\rangle\Big) \\ |1,-1\rangle = |--\rangle \end{cases} \tag{9.125}$$

et à $s = 0$

$$s = 0 \qquad |0,0\rangle = \frac{1}{\sqrt{2}}\Big(|+-\rangle - |-+\rangle\Big) \tag{9.126}$$

Comme nous avons trouvé quatre vecteurs orthogonaux, ces vecteurs forment une base de $\mathcal{H}_1 \otimes \mathcal{H}_2$ et l'ensemble des opérateurs compatibles $\{\vec{S}_1^2, \vec{S}_2^2, \vec{S}^{\,2}, S_z\}$, ou simplement $\{\vec{S}^{\,2}, S_z\}$, est bien complet. Les états $s = 1$ sont appelés *états triplets* et l'état $s = 0$, *état singulet.*

Comme application, diagonalisons l'opérateur $(\vec{\sigma}_1 \cdot \vec{\sigma}_2)$. Cet opérateur est diagonal dans la base $\{\vec{S}^{\,2}, S_z\}$. En effet,

$$\vec{S}^{\,2} = \frac{1}{4}(\vec{\sigma}_1 + \vec{\sigma}_2)^2 = \frac{3}{2} + \frac{1}{2}\,\vec{\sigma}_1 \cdot \vec{\sigma}_2 \tag{9.127}$$

ou

$$\vec{\sigma}_1 \cdot \vec{\sigma}_2 = 2\vec{S}^{\,2} - 3I = [2s(s+1) - 3]I$$

L'opérateur $\vec{\sigma}_1 \cdot \vec{\sigma}_2$ est égal à I dans l'état triplet et à $-3I$ dans l'état singulet. On en déduit les projecteurs $\mathcal{P}_1$ et $\mathcal{P}_0$ sur les états triplet et singulet

$$\mathcal{P}_0 + \mathcal{P}_1 = I \qquad \vec{\sigma}_1 \cdot \vec{\sigma}_2 = -3\,\mathcal{P}_0 + \mathcal{P}_1$$

d'où

$$\mathcal{P}_0 = \frac{1}{4}(I - \vec{\sigma}_1 \cdot \vec{\sigma}_2) \qquad \mathcal{P}_1 = \frac{1}{4}(3 + \vec{\sigma}_1 \cdot \vec{\sigma}_2) \tag{9.128}$$

L'opérateur $\vec{\sigma}_1 \cdot \vec{\sigma}_2$ est un opérateur scalaire qui commute avec $\vec{S}$, mais non avec les opérateurs de spin individuels $\vec{S}_1$ et $\vec{S}_2$. On remarque aussi que les états triplets sont symétriques (c'est-à-dire qu'ils ne changent pas de signe) dans la permutation des spins 1 et 2, tandis que l'état singulet est antisymétrique dans cette permutation.

9.6.2 Cas général : composition de deux moments angulaires $\vec{J}_1$ et $\vec{J}_2$

Nous allons généraliser ce qui précède à la composition de deux moments angulaires $\vec{J}_1$ et $\vec{J}_2$. Le raisonnement utilisé en (9.124) peut être répété pour montrer que $\vec{J} = \vec{J}_1 + \vec{J}_2$ est le moment angulaire total. Suivant la sous-section précédente, nous construisons l'espace produit tensoriel à $(2j_1+1)\times(2j_2+1)$ dimensions

$$\mathcal{E} = \mathcal{E}(j_1) \otimes \mathcal{E}(j_2)$$

Une base possible de cet espace est constituée des vecteurs propres

$$|j_1 j_2 m_1 m_2\rangle = |j_1 m_1\rangle \otimes |j_2 m_2\rangle \tag{9.129}$$

communs à $\vec{J}_1^2$, $\vec{J}_2^2$, J_{1z} et J_{2z}

$$\begin{aligned}
\vec{J}_1^2 |j_1 j_2 m_1 m_2\rangle &= j_1(j_1+1)\, |j_1 j_2 m_1 m_2\rangle \\
\vec{J}_2^2 |j_1 j_2 m_1 m_2\rangle &= j_2(j_2+1)\, |j_1 j_2 m_1 m_2\rangle \\
J_{1z} |j_1 j_2 m_1 m_2\rangle &= m_1\, |j_1 j_2 m_1 m_2\rangle \\
J_{2z} |j_1 j_2 m_1 m_2\rangle &= m_2\, |j_1 j_2 m_1 m_2\rangle
\end{aligned}$$

Cette base correspond à l'ensemble complet d'opérateurs compatibles $\{\vec{J}_1^2, \vec{J}_2^2, J_{1z}, J_{2z}\}$. Nous allons construire une autre base de $\mathcal{E}$, celle où les opérateurs $\{\vec{J}_1^2, \vec{J}_2^2, \vec{J}^{\,2}, J_z\}$ sont diagonaux. Nous partons des deux observations suivantes.

- Tout vecteur $|j_1 j_2 m_1 m_2\rangle$ est vecteur propre de J_z avec la valeur propre $m = m_1 + m_2$.
- Si une valeur de j est possible, on doit avoir par application de J_+ et J_- une série de $(2j+1)$ vecteurs $|jm\rangle$. *A priori*, on pourrait même avoir plusieurs séries de vecteurs de ce type, et nous noterons $N(j)$ le nombre de ces séries pour une valeur de j donnée.

Soit $n(m)$, la dégénérescence de la valeur propre m de J_z. Comme m apparaît si et seulement si $j \geq |m|$, on aura (figure 9.12)

$$n(m) = \sum_{j \geq |m|} N(j)$$

et par conséquent

$$N(j) = n(j) - n(j+1)$$

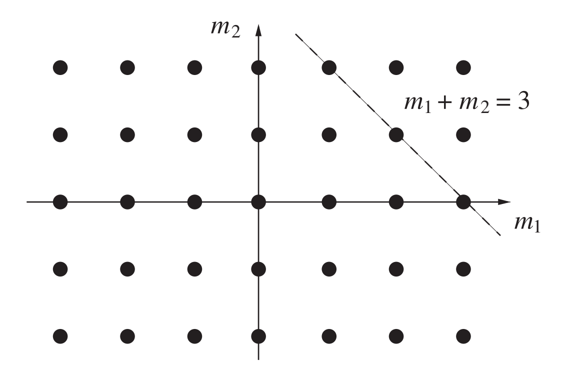


FIG. 9.12 – Composition de deux moments angulaires.

Mais $n(m)$ est égal au nombre de couples (m_1, m_2) tels que $m = m_1 + m_2$. En supposant par exemple $j_1 \geq j_2$

$$n(m) = \begin{cases} 0 & \text{si} & & |m| > j_1 + j_2 \\ j_1 + j_2 + 1 - |m| & \text{si} & j_1 - j_2 \leq & m \leq j_1 + j_2 \\ 2j_2 + 1 & \text{si} & 0 \leq & |m| \leq j_1 - j_2 \end{cases}$$

On en conclut

$$N(j) = 1 \quad \text{pour} \quad j_1 - j_2 \leq j \leq j_1 + j_2$$

et $N(j) = 0$ dans tous les autres cas. Pour tenir compte du cas $j_2 > j_1$, il suffit de remplacer $(j_1 - j_2)$ par $|j_1 - j_2|$. Nous pouvons donc énoncer le théorème suivant.

Théorème de composition des moments angulaires. Dans l'espace produit tensoriel $\mathcal{E} = \mathcal{E}(j_1) \otimes \mathcal{E}(j_2)$

1. Les valeurs possibles de j sont

$$|j_1 - j_2|, \; |j_1 - j_2| + 1, \; \cdots\cdots, j_1 + j_2 - 1, \; j_1 + j_2 \tag{9.130}$$

2. À chaque valeur de j correspond une seule série de vecteurs propres $|jm\rangle$

$$\vec{J}^{\,2}|jm\rangle = j(j+1)|jm\rangle \qquad J_z|jm\rangle = m|jm\rangle \tag{9.131}$$

Il est instructif de vérifier que la dimension de $\mathcal{E}$ est bien correcte ($j_1 \geq j_2$)

$$\begin{aligned}\dim \mathcal{E} &= \sum_{j_1 - j_2 \leq j \leq j_1 + j_2} (2j+1) \\ &= (j_1+j_2)(j_1+j_2+1) - (j_1-j_2-1)(j_1-j_2) \\ &\quad +(j_1+j_2) - (j_1-j_2-1) \\ &= (2j_1+1)(2j_2+1)\end{aligned}$$

On passe de la base orthonormée $|j_1 j_2 m_1 m_2\rangle$ à la base orthonormée $|jm\rangle$ par une transformation unitaire. Les éléments de la matrice unitaire qui effectue cette transformation sont appelés *coefficients de Clebsch-Gordan* (C-G) $C^{j_1 j_2}_{m_1 m_2;jm}$

$$|jm\rangle = \sum_{m_1+m_2=m} C^{j_1 j_2}_{m_1 m_2;jm} |j_1 j_2 m_1 m_2\rangle \tag{9.132}$$

qui ne peuvent être différents de zéro que si $m = m_1 + m_2$ et $|j_1 - j_2| \leq j \leq j_1 + j_2$. On choisit la convention de phase suivante

$$C^{j_1 j_2}_{m_1 m_2;j,m=j} \ \text{ réel} \geq 0$$

et on peut alors montrer par application de J_- que tous les C-G sont réels. Les coefficients de Clebsch-Gordan sont les éléments d'une matrice unitaire réelle, les indices matriciels étant $(m_1 m_2)$ et (jm). Ils vérifient donc les conditions d'orthogonalité

$$\sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} C^{j_1 j_2}_{m_1 m_2;jm} C^{j_1 j_2}_{m_1 m_2;j'm'} = \delta_{jj'} \delta_{mm'} \tag{9.133}$$

et inversement

$$\sum_{j=|j_1-j_2|}^{j_1+j_2} \sum_{m=-j}^{j} C^{j_1 j_2}_{m_1 m_2;jm} C^{j_1 j_2}_{m'_1 m'_2;jm} = \delta_{m_1 m'_1} \delta_{m_2 m'_2} \tag{9.134}$$

Les équations (9.125) et (9.126) donnent des exemples de C-G

$$C^{\frac{1}{2}\,\frac{1}{2}}_{\frac{1}{2}\,\frac{1}{2};11} = 1 \qquad C^{\frac{1}{2}\,\frac{1}{2}}_{\frac{1}{2}\,-\frac{1}{2};10} = \frac{1}{\sqrt{2}}$$

Comme application de la composition des moments angulaires, étudions le *couplage spin-orbite* : en raison d'effets relativistes, le moment angulaire orbital et le spin d'un électron atomique, par exemple l'électron de l'atome d'hydrogène ou l'électron de valence d'un alcalin, ne sont pas indépendants, comme nous le verrons au § 15.2.2. Le moment angulaire total de l'électron est la somme de son moment angulaire orbital $\vec{L}$ et de son spin $\vec{S}$

$$\vec{J} = \vec{L} + \vec{S} \tag{9.135}$$

Les valeurs possibles de j sont donc $j = l + 1/2$ et $j = l - 1/2$ (sauf si $l = 0$, auquel cas $j = s = 1/2$). Le moment angulaire orbital et le spin sont couplés par un potentiel spin-orbite

$$V_{\rm so}(r) = V(r)\,\vec{L}\cdot\vec{S} \tag{9.136}$$

Ce potentiel prend des valeurs différentes selon que $j = l+1/2$ ou $j = l-1/2$; en effet,

$$(\vec{L}+\vec{S})^2 = \vec{J}^{\,2} = \vec{L}^{\,2} + \vec{S}^{\,2} + 2\vec{L}\cdot\vec{S}$$

soit

$$\vec{L}\cdot\vec{S} = \frac{1}{2}\Big[j(j+1) - l(l+1) - s(s+1)\Big] \tag{9.137}$$

ce qui donne pour le potentiel spin-orbite

$$V_{\rm so}(r) = \begin{cases} \frac{1}{2}\,V(r)l & : \quad j = l+1/2 \\ -\frac{1}{2}\,V(r)(l+1) & : \quad j = l-1/2 \end{cases} \tag{9.138}$$

Les coefficients de Clebsch-Gordan du couplage spin-orbite sont calculés dans l'exercice 9.7.18.

9.6.3 Composition des matrices de rotation

La loi de composition des moments angulaires se reflète dans une loi de composition des matrices de rotation. Considérons les éléments de matrice de l'opérateur de rotation $U(\mathcal{R})$ entre les états $|jm\rangle$ et $|jm'\rangle$ du type (9.132)

$$\begin{aligned} \langle jm|U(\mathcal{R})|jm'\rangle &= D^{(j)}_{mm'}(\mathcal{R}) \\ &= \sum_{m_1m_2}\sum_{m'_1m'_2} C^{j_1j_2}_{m_1m_2;jm}C^{j_1j_2}_{m'_1m'_2;jm'} \\ &\quad\ \langle j_1j_2m_1m_2|U(\mathcal{R})|j_1j_2m'_1m'_2\rangle \end{aligned}$$

d'où

$$D^{(j)}_{mm'}(\mathcal{R}) = \sum_{m_1m_2}\sum_{m'_1m'_2} C^{j_1j_2}_{m_1m_2;jm}C^{j_1j_2}_{m'_1m'_2;jm'}D^{(j_1)}_{m_1m'_1}(\mathcal{R})D^{(j_2)}_{m_2m'_2}(\mathcal{R}) \tag{9.139}$$

Compte tenu des relations d'orthogonalité (9.133) et (9.134) des C-G, on peut inverser (9.139)

$$D^{(j_1)}_{m_1m'_1}(\mathcal{R})D^{(j_2)}_{m_2m'_2}(\mathcal{R}) = \sum_{|j_1-j_2|}^{j_1+j_2} C^{j_1j_2}_{m_1m_2;jm}C^{j_1j_2}_{m'_1m'_2;jm'}D^{(j)}_{mm'}(\mathcal{R}) \tag{9.140}$$

Cette équation peut être interprétée de la manière suivante : dans l'espace $\mathcal{E}(j_1)\otimes\mathcal{E}(j_2)$, on forme la matrice $\Delta(\mathcal{R})$, produit tensoriel de $D^{(j_1)}(\mathcal{R})$ et $D^{(j_2)}(\mathcal{R})$

$$\Delta_{m_1m_2;m'_1m'_2}(\mathcal{R}) = D^{(j_1)}_{m_1m'_1}(\mathcal{R})\otimes D^{(j_2)}_{m_2m'_2}(\mathcal{R})$$

Par un changement de base effectué grâce à une matrice unitaire dont les éléments sont les coefficients de C-G $C^{j_1 j_2}_{m_1 m_2;jm}$, la matrice

$$\Delta'(\mathcal{R}) = C\Delta(\mathcal{R})C^{-1}$$

devient une matrice diagonale par blocs

$$C\Delta(\mathcal{R})C^{-1} = \begin{pmatrix} D^{(j_1+j_2)} & 0 & \dots & 0 \\ 0 & D^{(j_1+j_2-1)} & \dots & \vdots \\ 0 & 0 & \dots & \vdots \\ 0 & \dots & 0 & D^{(|j_1-j_2|} \end{pmatrix}$$

En termes mathématiques, ceci s'appelle réduire le produit de deux représentations $D^{(j_1)}$ et $D^{(j_2)}$ du groupe de rotation en composantes irréductibles

$$D^{(j_1)} \otimes D^{(j_2)} = D^{(j_1+j_2)} + D^{(j_1-j_2-1)} + \dots + D^{(|j_1-j_2|)} \tag{9.141}$$

9.6.4 Théorème de Wigner-Eckart (opérateurs scalaires et vectoriels)

Nous avons donné en (7.33) la définition d'un opérateur scalaire $\mathcal{S}$: c'est un opérateur qui commute avec $\vec{J}$: $[\mathcal{S}, \vec{J}] = 0$. Examinons les éléments de matrice $\langle j'm'|\mathcal{S}|jm\rangle$ de $\mathcal{S}$ dans une base standard du moment angulaire

$$[\mathcal{S}, \vec{J}^2] = 0 \Rightarrow j' = j \qquad [\mathcal{S}, J_z] = 0 \Rightarrow m' = m$$

De plus,

$$[\mathcal{S}, J_\pm] = 0 \Rightarrow \langle jm|\mathcal{S}|jm\rangle = \langle j||\mathcal{S}||j\rangle \ \text{ indépendant de } m \tag{9.142}$$

La quantité $\langle j||\mathcal{S}||j\rangle$ est appelée *élément de matrice réduit* de $\mathcal{S}$.

Nous passons maintenant aux opérateurs vectoriels $\vec{V}$, dont nous avons donné la définition en (7.34) : les composantes cartésiennes V_k d'un opérateur vectoriel se transforment par rotation suivant la loi

$$U^\dagger(\mathcal{R})\, V_k\, U(\mathcal{R}) = \sum_l \mathcal{R}_{kl}\, V_l \tag{9.143}$$

En considérant des rotations infinitésimales, nous en avions déduit les relations de commutation avec les composantes du moment angulaire

$$[J_k, V_l] = \mathrm{i} \sum_p \varepsilon_{klp} V_p \tag{9.144}$$

Les équations (9.143) et (9.144) sont strictement équivalentes et peuvent servir l'une comme l'autre pour définir un opérateur vectoriel. Il est commode d'utiliser les composantes sphériques V_q de $\vec{V}$

$$V_1 = -\frac{1}{\sqrt{2}}\,(V_x + \mathrm{i}V_y) \qquad V_0 = V_z \qquad V_{-1} = \frac{1}{\sqrt{2}}\,(V_x - \mathrm{i}V_y) \tag{9.145}$$

Ces composantes sont aussi appelées composantes standard de $\vec{V}$, car dans le cas où $\vec{V}$ est l'opérateur position, $\vec{V} = \vec{R}$, les composantes $\hat{r}_1, \hat{r}_0$ et $\hat{r}_{-1}$ du vecteur $\hat{r}$ ne sont autres que les harmoniques sphériques $Y_1^{\pm}$ et Y_1^0 à un facteur $\sqrt{3/4\pi}$ près (*cf.* (9.64)). Ceci implique d'après (9.65) la loi de transformation

$$(\mathcal{R}\hat{r})_m = \sum_{m'} D^{(1)}_{m'm}(\mathcal{R}^{-1})\hat{r}_{m'} \tag{9.146}$$

La loi de transformation des composantes sphériques de $\vec{V}$ est donc[14]

$$U(\mathcal{R})\, V_q\, U^\dagger(\mathcal{R}) = \sum_{q'} D^{(1)}_{q'q}(\mathcal{R}) V_{q'} \tag{9.147}$$

On pourrait naturellement vérifier cette expression en utilisant les formules explicites pour $D^{(1)}$ et la définition des composantes sphériques.

Notre objectif est de relier les éléments de matrice des diverses composantes d'un opérateur vectoriel entre des états $|jm\rangle$. Pour ce faire, examinons les propriétés de transformation par rotation du vecteur $|1jqm\rangle = V_q|jm\rangle$

$$\begin{aligned} U(\mathcal{R})|1jqm\rangle &= U(\mathcal{R})V_q U^\dagger(\mathcal{R})U(\mathcal{R})|jm\rangle \\ &= \sum_{q'm'} D^{(1)}_{q'q}(\mathcal{R})D^{(j)}_{m'm}(\mathcal{R})|1jq'm'\rangle \end{aligned}$$

Les vecteurs $|1jqm\rangle$ se transforment par rotation exactement de la même façon que les vecteurs $|j_1 j_2 m_1 m_2\rangle$ avec $j_1 = 1, j_2 = j, m_1 = q, m_2 = m$. On peut donc construire des vecteurs

$$|\tilde{\jmath}\,\tilde{m}\rangle = \sum_{m+q=\tilde{m}} C^{1j}_{qm;\tilde{\jmath}\,\tilde{m}}\, |1jqm\rangle \tag{9.148}$$

qui se transforment par rotation suivant la loi

$$U(\mathcal{R})|\tilde{\jmath}\,\tilde{m}\rangle = \sum_{\tilde{m}'} D^{(\tilde{\jmath})}_{\tilde{m}'\tilde{m}}|\tilde{\jmath}\,\tilde{m}'\rangle$$

Cette équation montre que les vecteurs $|\tilde{\jmath}\,\tilde{m}\rangle$ forment une base standard de l'espace $\mathcal{E}(\tilde{\jmath})$, à un facteur global multiplicatif près. Ces vecteurs ne seront pas en général normalisés à un, mais ils auront une norme identique, quel que soit $\tilde{m}$

$$\langle\tilde{\jmath}\,\tilde{m}|\tilde{\jmath}'\tilde{m}'\rangle = \delta_{\tilde{\jmath}\tilde{\jmath}'}\delta_{\tilde{m}\tilde{m}'}\alpha(\tilde{\jmath})$$

Inversant (9.148)

$$V_q|jm\rangle = |1qjm\rangle = \sum_{\tilde{\jmath}=j-1}^{j+1} C^{1j}_{qm;\tilde{\jmath}\,\tilde{m}}|\tilde{\jmath}\,\tilde{m}\rangle$$

14. On remarquera que l'ordre U, $U^\dagger$, ainsi que l'ordre des indices, sont différents de ceux de (9.143).

d'où

$$\begin{aligned}\langle j'm'|V_q|jm\rangle &= \sum_{\tilde{j}} C^{1j}_{qm;\tilde{j}\,\tilde{m}}\,\langle j'm'|1j\tilde{j}\,\tilde{m}\rangle \\ &= \sum_{\tilde{j}} C^{1j}_{qm;\tilde{j}\,\tilde{m}}\,\delta_{j'\tilde{j}}\delta_{m'\tilde{m}}\beta(j',j) = C^{1j}_{qm;j',m'}\,\beta(j',j)\end{aligned}$$

Définissant l'élément de matrice réduit $\langle j'||V_q||j\rangle$ par

$$\langle j'||V||j\rangle = \beta(j',j)$$

on obtient le *théorème de Wigner-Eckart* pour les opérateurs vectoriels

$$\langle j'm'|V_q|jm\rangle = C^{1j}_{qm;j'm'}\,\langle j'||V||j\rangle \qquad (9.149)$$

Toute la dépendance dans les nombres quantiques magnétiques m, m' et q est contenue dans le coefficient de Clebsch-Gordan $C^{1j}_{qm;j'm'}$, que l'on trouve dans des tables. À j fixé, les seules valeurs possibles de j' sont $j' = j-1,\ j,\ j+1$. Ce théorème se généralise aux *opérateurs tensoriels irréductibles* : voir l'exercice 9.7.19.

Comme application, calculons les éléments de matrice d'un opérateur vectoriel lorsque $j = j'$, en utilisant le fait que $\vec{J}$ est un opérateur vectoriel dont les éléments de matrice obéissent à (9.149)

$$\langle jm'|J_q|jm\rangle = C^{1j}_{qm;jm'}\langle j||J||j\rangle$$

Il en découle une relation de proportionnalité pour les composantes cartésiennes V_k

$$\langle jm'|V_k|jm\rangle = K\,\langle jm'|J_k|jm\rangle$$

Pour évaluer la constante K, nous calculons le produit scalaire $\vec{J}\cdot\vec{V}$, qui est un opérateur scalaire

$$\begin{aligned}\langle jm|(\vec{J}\cdot\vec{V})|jm\rangle &= \sum_{k,m''}\langle jm|J_k|jm''\rangle\langle jm''|V_k|jm\rangle \\ &= K\sum_{k,m''}\langle jm|J_k|jm''\rangle\langle jm''|J_k|jm\rangle \\ &= K\,\langle jm|\vec{J}^{\,2}|jm\rangle = K\,j(j+1)\end{aligned}$$

En combinant ces équations, on obtient pour les éléments de matrice de V_k

$$\langle jm'|V_k|jm\rangle = \frac{1}{j(j+1)}\,\langle j||(\vec{J}\cdot\vec{V})||j\rangle\langle jm'|J_k|jm\rangle \qquad (9.150)$$

Comme $(\vec{J}\cdot\vec{V})$ est un opérateur scalaire, $\langle jm|(\vec{J}\cdot\vec{V})|jm\rangle$ est indépendant de m et égal à l'élément de matrice réduit $\langle j||(\vec{J}\cdot\vec{V})||j\rangle$.

9.7 Exercices

9.7.1 Propriétés de $\vec{J}$

Vérifier par un calcul explicite que $[\vec{J}^2, J_z] = 0$. Vérifier également les identités (9.5) à (9.9).

9.7.2 Rotation d'un moment angulaire

Soit $\mathcal{R}$ la rotation (9.30) d'angles (θ, ϕ). Vérifier que le vecteur

$$U(\mathcal{R})|jm\rangle = \mathrm{e}^{-\mathrm{i}\phi J_z}\, \mathrm{e}^{-\mathrm{i}\theta J_y}|jm\rangle$$

est vecteur propre de l'opérateur

$$J_x \sin\theta \cos\phi + J_y \sin\theta \sin\phi + J_z \cos\theta = \vec{J} \cdot \hat{n}$$

avec la valeur propre m ; $\hat{n}$ est le vecteur unitaire dans la direction (θ, ϕ). Suggestion : adapter (7.29).

9.7.3 Rotations (θ, ϕ)

Montrer que l'on peut écrire la rotation (9.30) $\mathcal{R}(\theta, \phi)$ sous la forme

$$\mathcal{R}(\theta, \phi) = \mathcal{R}_{y'}(\theta)\mathcal{R}_z(\phi)$$

où Oy' est l'axe obtenu à partir de Oy par une rotation de ϕ autour de Oz. Suggestion : montrer

$$\mathcal{R}_{y'}(\theta) = \mathcal{R}_z(\phi)\mathcal{R}_y(\theta)\mathcal{R}_z(-\phi)$$

9.7.4 Moments angulaires $j = \frac{1}{2}$ et $j = 1$

1. Retrouver à partir de (9.23) les opérateurs S_x, S_y et S_z du spin 1/2.

2. Toujours à partir de (9.23), calculer les matrices 3×3 représentatives de J_x, J_y et J_z pour le moment angulaire $j = 1$.

3. Montrer que pour $j = 1$, J_x, J_y, J_z sont reliés aux générateurs infinitésimaux (7.26) T_x, T_y, T_z par la transformation unitaire qui fait passer des composantes cartésiennes de $\hat{r}$ à ses composantes sphériques (9.64) : $J_i = U^\dagger T_i U$ avec

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 0 & 1 \\ -\mathrm{i} & 0 & -\mathrm{i} \\ 0 & \sqrt{2} & 0 \end{pmatrix}$$

4. Calculer la matrice de rotation $d^{(1)}(\theta)$

$$d^{(1)}(\theta) = \exp(-\mathrm{i}\theta J_y)$$

et vérifier (9.39). Suggestion : montrer que $J_y^3 = J_y$.

9.7.5 Moment angulaire orbital

1. Utiliser les relations de commutation canoniques

$$[X_i, P_j] = \mathrm{i}\, \delta_{ij} I$$

et l'expression $\vec{L} = \vec{R} \times \vec{P}$ pour montrer

$$[L_x, L_y] = \mathrm{i} L_z$$

2. Démontrer les équations (9.47) à (9.49). Suggestion : montrer que pour une rotation infinitésimale d'angle $\mathrm{d}\alpha$ autour de Ox, les angles θ et ϕ varient de

$$\mathrm{d}\theta = -\sin\phi\, \mathrm{d}\alpha \qquad \mathrm{d}\phi = -\frac{\cos\phi}{\tan\theta}\, \mathrm{d}\alpha$$

En déduire L_x et $L_y = \mathrm{i}[L_x, L_z]$.

3. Comme $L_z = -\mathrm{i}\partial/\partial\phi$, on pourrait s'attendre à une inégalité de Heisenberg

$$\Delta\phi\, \Delta L_z \geq \frac{1}{2}$$

Or, dans un état propre de L_z où m est fixé, $\Delta L_z = 0$, tandis que $\Delta\phi \leq 2\pi$, puisque $0 \leq \phi \leq 2\pi$ et l'inégalité de Heisenberg est violée dans cet état. Où est la faute de raisonnement ? Suggestion : voir l'exercice 6.4.3, question 2. Pourquoi le raisonnement de l'exercice 8.6.1 est-il en défaut ?

9.7.6 Relation entre les matrices de rotation et les harmoniques sphériques

1. Soit $\varphi(\vec{r}) = \varphi(x, y, z)$, la fonction d'onde d'une particule. Montrer que

$$\left(\mathrm{e}^{-\mathrm{i}\alpha L_z} \varphi\right)(0, 0, z) = \varphi(0, 0, z)$$

et en déduire que si une particule est localisée sur l'axe Oz, la composante z de son moment angulaire orbital est nulle. Interpréter qualitativement ce résultat.

2. On suppose que le moment angulaire orbital de la particule est l et on écrit sa fonction d'onde comme le produit d'un harmonique sphérique et d'une fonction d'onde radiale $g_l(r)$ qui ne dépend que de $r = |\vec{r}|$

$$\psi_{lm}(\vec{r}) = Y_l^m(\theta, \phi) g_l(r) = \langle \theta, \phi | lm \rangle g_l(r)$$

On s'intéresse uniquement à la partie angulaire. En utilisant

$$|\theta, \phi\rangle = U(\mathcal{R}) |\theta = 0, \phi = 0\rangle$$

où $\mathcal{R}$ est la rotation d'angles (θ, ϕ), montrer que

$$Y_l^m(\theta, \phi) \propto \left[D_{m0}^{(l)}(\theta, \phi) \right]^*$$

On peut montrer que le coefficient de proportionnalité est $\sqrt{(2l+1)/(4\pi)}$

$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}}\left[D^{(l)}_{m0}(\theta,\phi)\right]^*$$

9.7.7 Indépendance de l'énergie par rapport à m

En supposant le potentiel $V(r)$ invariant par rotation, soit ψ_{lm} une solution de l'équation de Schrödinger indépendante du temps

$$H\psi_{lm} = E_{lm}\psi_{lm}$$

Utiliser la relation de commutation $[L_+, H] = 0$ pour montrer que l'énergie E_{lm} est en fait indépendante de m.

9.7.8 Puits sphérique

1. Soit le potentiel $V(\vec{r})$ à symétrie sphérique (voir la figure 13.4)

$$\begin{aligned} V(\vec{r}) &= -V_0 \qquad 0 \le r \le R \\ &= 0 \qquad\qquad r > R \end{aligned}$$

appelé puits sphérique. Établir l'équation donnant les états liés dans l'onde s ($l = 0$). Y a-t-il toujours un état lié ? Comparer avec le cas du puits à une dimension.

2. On modélise le potentiel neutron-proton par un puits sphérique de rayon $R \simeq 2\,\text{fm}$. Il existe un seul état lié neutron-proton dans l'onde s, le deutéron[15], dont l'énergie de liaison est $B \simeq 2.2\,\text{MeV}$. Calculer la profondeur V_0 du puits nécessaire pour qu'il y ait juste un état lié. Comparer V_0 avec l'énergie de liaison et montrer que $V_0 \gg B$.

3. Trouver les niveaux d'énergie dans l'onde s d'une particule dans le potentiel

$$V(r) = \frac{A}{r^2} - \frac{B}{r} \qquad A, B > 0$$

9.7.9 Atome d'hydrogène pour $l \neq 0$

1. Écrire l'équation qui généralise (7.26) lorsque le moment angulaire orbital $l \neq 0$. En déduire que l'on doit rajouter à (9.91) le terme

$$-l(l+1)\left[\frac{a_1}{x} + \sum_{k=1}^{\infty} a_{k+1}x^{k-1}\right]$$

2. Montrer la relation de récurrence

$$a_{k+1} = \frac{2(\alpha k - 1)}{k(k+1) - l(l+1)}\, a_k$$

15. En fait, le deutéron a ausssi une petite composante d'onde d.

et en déduire $\alpha = \dfrac{1}{n}$, $k \geq l+1$ et par conséquent $l+1 \leq k \leq n$. Montrer que le spectre de l'atome d'hydrogène est donné par (9.93).

9.7.10 Éléments de matrice d'un potentiel

L'électron externe d'un atome est supposé dans un état p ($l = 1$). Sa fonction d'onde est

$$\psi_{1m}(\vec{r}) = Y_1^m(\theta, \phi)\, \frac{u_1(r)}{r}$$

Il est plongé dans un potentiel extérieur de la forme

$$V(\vec{r}) = Ax^2 + By^2 - (A+B)z^2$$

où A et B sont des constantes.

1. Montrer *sans calculs* que la matrice représentative de V dans la base $|lm\rangle$ est de la forme

$$V_{m'm} = \begin{pmatrix} \alpha & 0 & \beta \\ 0 & \gamma & 0 \\ \beta & 0 & \alpha \end{pmatrix}$$

les lignes et les colonnes étant rangées dans l'ordre m', $m = 1, 0, -1$.

2. Déterminer les valeurs propres et vecteurs propres de V. Montrer que $\langle L_z \rangle = 0$ dans un état propre de V.

3. Utiliser (9.63) pour calculer explicitement α, β et γ en fonction de A, B et

$$I = \int_0^\infty |u_1(r)|^2 r^2 dr$$

.

9.7.11 Équation radiale en dimension $d = 2$

On se propose d'écrire l'équivalent de l'équation (9.78) lorsque l'espace est à deux dimensions et le potentiel invariant par rotation. L'équation de Schrödinger indépendante du temps est

$$\left[-\frac{1}{2M}\nabla^2 + V(r)\right]\psi(\vec{r}) = E\psi(\vec{r})$$

On utilise les coordonnées polaires dans le plan xOy

$$x = r\cos\theta \qquad y = r\sin\theta$$

On rappelle l'expression du laplacien en coordonnées polaires

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}$$

et l'expression du moment angulaire

$$L_z = XP_y - YP_x = -\mathrm{i}\,\frac{\partial}{\partial\theta}$$

1. Montrer que les fonctions propres de L_z sont de la forme $\exp(\mathrm{i}m\theta)$.

2. On cherche des solutions de l'équation de Schrödinger de la forme

$$\psi_{nm}(\vec{r}) = \frac{1}{\sqrt{r}}\,\mathrm{e}^{\,\mathrm{i}m\theta}u_{nm}(r)$$

Montrer que $u_{nm}(r)$ et E_{nm} vérifient l'équation radiale

$$\left[-\frac{1}{2M}\,\frac{\mathrm{d}^2}{\mathrm{d}r^2} + V(r) + \frac{m^2 - 1/4}{2Mr^2}\right]u_{nm}(r) = Eu_{nm}(r)$$

Quelle est l'interprétation de n ? Quel est le comportement de $u_{nm}(r)$ lorsque $r \to 0$?

9.7.12 Propriété de symétrie des matrices $d^{(j)}$

En utilisant l'opérateur Y (9.100), démontrer la propriété de symétrie des matrices de rotation $d^{(j)}(\beta)$

$$d^{(j)}_{m'm}(\beta) = (-1)^{m-m'}\,d^{(j)}_{-m',-m}(-\beta)$$

9.7.13 Diffusion de la lumière

1. On reprend l'étude de la transition radiative $A^* \to A + \gamma$ avec $j = j(A^*) = 1$ et $j' = j(A) = 0$. Déterminer dans le cas dipolaire électrique les amplitudes de transition pour un état initial $m = 1$ lorsque les photons polarisés circulairement sont émis dans le plan xOz avec une impulsion $\vec{p}$ faisant un angle θ avec l'axe Oz

$$\begin{aligned} a_D^{m=1}(\theta) &= \frac{1}{2}\,a(1+\cos\theta) \\ a_G^{m=1}(\theta) &= \frac{1}{2}\,a(1-\cos\theta) \end{aligned}$$

Généraliser au cas où le photon est émis dans la direction (θ,ϕ).

2. On suppose que des photons d'impulsion $\vec{p} \parallel Oz$ arrivent sur l'atome dans son état fondamental A. L'atome absorbe le photon et est porté dans son état excité A^* ; il revient dans son état fondamental en émettant un photon dans une direction du plan xOz faisant un angle θ avec Oz. On appelle b l'amplitude d'absorption d'un photon de polarisation circulaire droite D

$$b = \langle j=1, m=1|T'|D\rangle$$

Montrer que si les transitions sont dipolaires électriques, on a aussi

$$b = \langle j = 1, m = -1 | T' | G \rangle$$

Soit $c_{P \to P'}(\theta)$ l'amplitude de transition pour la diffusion d'un photon initial de polarisation circulaire P ($P = D$ ou G) sous un angle θ avec une polarisation finale P'. Montrer que

$$c_{P \to P'}(\theta) = \frac{ab}{2}\,(1 \pm \cos\theta)$$

où le signe (+) correspond à $P = P'$ et le signe (−) à $P \neq P'$. En déduire pour une polarisation linéaire $|x\rangle$ du photon initial et pour des polarisations linéaires $|x'\rangle$ ou $|y\rangle$ du photon diffusé, définies par rapport à la direction de propagation de ce photon

$$\begin{aligned} c_{x \to x'}(\theta) &= ab\cos\theta \\ c_{x \to y}(\theta) &= 0 \end{aligned}$$

Donner une analogie classique qui conduit également à une distribution angulaire en $\cos^2\theta$ avec un rayonnement polarisé dans le plan xOz. Généraliser au cas où le photon est émis dans la direction (θ, ϕ).

9.7.14 Mesure du moment magnétique du Λ^0

La particule Λ^0 est une particule de charge nulle, de masse $M \simeq 1115\,\mathrm{MeV/c^2}$, de spin 1/2 et de vie moyenne $\tau \simeq 2.5 \times 10^{-10}\,\mathrm{s}$. Un de ses modes de désintégration principaux (66 % des cas) est

$$\Lambda^0 \to \text{proton} + \text{méson } \pi^-$$

où le proton a un spin 1/2 et le méson π^- un spin 0.

1. Dans le référentiel où le Λ^0 est au repos, on suppose le proton émis avec une impulsion $\vec{p}$ dans la direction Oz, choisie comme axe de quantification du moment angulaire. Soit m, la projection suivant Oz du spin du Λ^0 et m' celle du proton. Pourquoi doit-on avoir $m = m'$? Soit a et b les amplitudes de probabilité des transitions

$$\begin{aligned} a \quad &: \quad \Lambda^0\left(m = \frac{1}{2}\right) \to \text{proton} \ \left(m' = \frac{1}{2};\ \vec{p} \parallel Oz\right) \\ b \quad &: \quad \Lambda^0\left(m = -\frac{1}{2}\right) \to \text{proton} \ \left(m' = -\frac{1}{2};\ \vec{p} \parallel Oz\right) \end{aligned}$$

Montrer que $|a| = |b|$ si la parité est conservée dans la désintégration. Suggestion : examiner l'action d'une réflexion par rapport au plan xOz.

2. Le proton est maintenant émis avec une impulsion $\vec{p}$ parallèle à une direction $\hat{n}$ du plan xOz, faisant un angle θ avec Oz. Soit m', la projection du spin du proton *sur la direction* $\hat{n}$ et $a_{m'm}(\theta)$ l'amplitude

$$a_{m'm}(\theta) \ : \ \Lambda^0 \left(m = \frac{1}{2}\right) \to \text{proton} \ \ (m'; \ \vec{p} \parallel \hat{n})$$

Exprimer

$$a_{\frac{1}{2},\frac{1}{2}}(\theta) = a_{++}(\theta) \ \text{ et } \ a_{-\frac{1}{2},\frac{1}{2}}(\theta) = a_{-+}(\theta)$$

en fonction de a, b et θ.

3. On suppose que le Λ^0 est produit dans un état de spin $m = 1/2$. Montrer que la distribution angulaire du proton est de la forme

$$w(\theta) = w_0(1 + \alpha \cos\theta)$$

Calculer α en fonction de a et b. L'expérience montre que

$$\alpha \simeq -0.645 \pm 0.016$$

Que peut-on en conclure sur la conservation de la parité dans la désintégration ?

4. Le Λ^0 est produit en bombardant une cible de protons au repos par un faisceau de mésons π^- dans la réaction (figure 9.13)

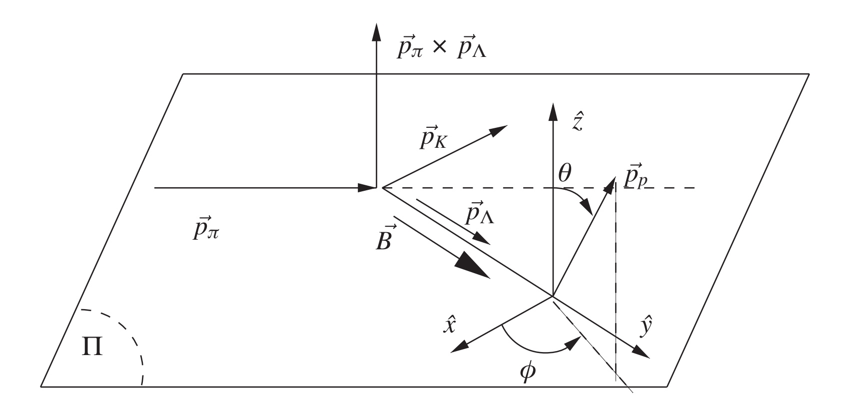


FIG. 9.13 – Cinématique de la production et de la désintégration du Λ^0.

$$\text{méson } \pi^- + \text{proton} \to \Lambda^0 + \text{méson } K^0$$

Par conservation de l'impulsion, $\vec{p}_{\pi^-}$, $\vec{p}_{\Lambda^0}$ et $\vec{p}_{K^0}$ sont situés dans un même plan π. On choisit pour axe Oz une perpendiculaire à ce plan

$$\hat{z} = \frac{\vec{p}_{\pi^-} \times \vec{p}_{\Lambda^0}}{|\vec{p}_{\pi^-} \times \vec{p}_{\Lambda^0}|}$$

et comme axe Oy la direction $\vec{p}_{\Lambda^0}$ de l'impulsion du Λ^0. Sachant que la parité est conservée dans la réaction de production et que les protons cibles ne sont pas polarisés, montrer que si $\vec{S}$ est l'opérateur de spin du Λ^0, alors les valeurs moyennes des composantes S_x et S_y sont nulles : $\langle S_x \rangle = \langle S_y \rangle = 0$.

5. On suppose pour simplifier[16] que $\langle S_z \rangle = 1/2$ et que tous les Λ^0 ont la même durée de vie τ et se désintègrent au même point. Le système est plongé dans un champ magnétique $\vec{B}$ uniforme et constant parallèle à Oy. Le Λ^0 possède un moment magnétique $\vec{\mu}$ relié à son spin $\vec{S}$ par un facteur gyromagnétique γ : $\vec{\mu} = \gamma \vec{S}$. Décrire qualitativement le mouvement du spin du Λ^0. Déterminer son orientation au moment de la désintégration en fonction de τ, de B et γ. Montrer que la distribution angulaire du proton émis dans la désintégration est

$$w(\theta, \phi) = w_0(1 + \alpha \cos \Theta)$$

avec

$$\cos \Theta = \cos \lambda \cos \theta + \sin \lambda \sin \theta \cos \phi$$

où les angles θ et ϕ sont les angles polaire et azimutal de l'impulsion du proton. Quelle est la valeur de l'angle λ ? En déduire que la détermination de $w(\theta, \phi)$ permet de mesurer le facteur gyromagnétique γ. On négligera la courbure de la trajectoire du proton due au champ magnétique ainsi que les transformations des angles dues au mouvement du Λ^0.

9.7.15 Production et désintégration du méson ρ^+

1. Le méson ρ^+ est une particule de spin 1 qui se désintègre en deux mésons π, particules de spin 0

$$\rho^+ \to \pi^+ + \pi^0$$

On choisit de se placer dans un référentiel où le méson ρ^+ est au repos ; on suppose que son spin est quantifié suivant l'axe Oz et qu'il se trouve initialement dans l'état de spin $|1m\rangle$, $m = -1, 0, 1$. Soit

$$a_m(\theta, \phi) = \langle \theta, \phi | T | 1m \rangle$$

l'amplitude de transition pour la désintégration du méson ρ^+ dans l'état initial $|1m\rangle$ avec émision du méson π^+ dans la direction d'angles polaire et azimutal (θ, ϕ). Montrer que l'on peut écrire

$$a_m(\theta, \phi) = a \left[D^{(1)}_{m0}(\theta, \phi) \right]^*$$

Quelle est la signification physique de a ? En déduire la distribution angulaire $W_m(\theta, \phi)$ du méson π^+, c'est-à-dire la probabilité d'émission du méson π^+

16. En fait, $|\langle S_z \rangle| < 1/2$ et on doit faire appel au formalisme de l'opérateur statistique : section 11.1.

dans la direction (θ, ϕ) lorsque le méson ρ^+ est initialement dans l'état $|1m\rangle$. Montrer que $W_m(\theta, \phi)$ est indépendant de ϕ (pourquoi ?) et donner son expression explicite en fonction de θ pour les trois valeurs de m, $m = -1, 0, 1$.

2. Si l'état initial du méson ρ^+ est une combinaison linéaire d'états $|1m\rangle$

$$|\lambda\rangle = \sum_{m=-1,0,1} c_m |1m\rangle \qquad \sum_{m=-1,0,1} |c_m|^2 = 1$$

quelle sera la distribution angulaire $W_\lambda(\theta, \phi)$?

3. En général, le méson ρ^+ n'est pas produit dans un état pur, mais dans un mélange décrit par un opérateur statistique ρ (section 11.1)

$$\rho = \sum_\lambda \mathsf{p}_\lambda |\lambda\rangle\langle\lambda| \qquad \mathsf{p}_\lambda \geq 0 \qquad \sum_\lambda \mathsf{p}_\lambda = 1$$

Montrer que la distribution angulaire est alors

$$\begin{aligned} W(\theta, \phi) &= \rho_{00} \cos^2\theta + \frac{1}{2} \sin^2\theta(\rho_{11} + \rho_{-1,-1}) \\ &+ \frac{1}{\sqrt{2}} \sin 2\theta \, \mathrm{Re}\left(\rho_{-10}\, \mathrm{e}^{-\mathrm{i}\phi} - \rho_{10}\, \mathrm{e}^{\mathrm{i}\phi}\right) - \sin^2\theta \, \mathrm{Re}\left(\rho_{1,-1}\, \mathrm{e}^{2\mathrm{i}\phi}\right) \end{aligned}$$

4. Le méson ρ^+ est produit dans la réaction : méson π^+ $(\vec{p}_1)$+proton$(\vec{p} = 0)$ $\to$ méson $\rho^+(\vec{p}_2)$+proton$(\vec{p}_3)$, où $\vec{p}_i$ dénote l'impulsion des particules. On choisit pour axe Oz la normale $\hat{n}$ au plan de la réaction

$$\hat{n} = \frac{\vec{p}_1 \times \vec{p}_2}{|\vec{p}_1 \times \vec{p}_2|}$$

La parité Π est conservée dans cette réaction et on suppose que les protons cibles ne sont pas polarisés. Montrer que la valeur moyenne $\langle \vec{J} \rangle = \mathrm{Tr}(\rho \vec{J})$ du spin du méson ρ^+ est dirigée suivant $\hat{n}$: $\langle \vec{J} \rangle = c\hat{n}$. En déduire

$$\mathrm{Tr}\,(\rho J_x) = \mathrm{Tr}\,(\rho J_y) = 0$$

Utiliser le fait que la cinématique de la réaction de production est invariante dans l'opération

$$\mathcal{Z} = \Pi\, \mathrm{e}^{-\mathrm{i}\pi J_z} \qquad [\rho, \mathcal{Z}] = 0$$

pour montrer que

$$\rho_{mm'} = (-1)^{m-m'} \rho_{mm'}$$

ρ ne dépend, en fait, que de quatre paramètres *réels* et a la forme d'un damier

$$\begin{pmatrix} \rho_{11} & 0 & \rho_{1,-1} \\ 0 & \rho_{00} & 0 \\ \rho^*_{1,-1} & 0 & \rho_{-1,-1} \end{pmatrix}$$

9.7.16 Interaction de deux dipôles

Le hamiltonien d'interaction de deux dipôles magnétiques portés par des particules de spin 1/2 s'écrit

$$H = \frac{K}{r^3}\left[3(\vec{\sigma}_1\cdot\hat{r})(\vec{\sigma}_2\cdot\hat{r}) - \vec{\sigma}_1\cdot\vec{\sigma}_2\right] = \frac{K}{r^3}\,S_{12}$$

où $\vec{r}$ est le vecteur joignant les deux dipôles et $\vec{\sigma}_1$ et $\vec{\sigma}_2$ les matrices de Pauli des deux particules. Soit

$$\vec{\Sigma} = \frac{1}{2}(\vec{\sigma}_1 + \vec{\sigma}_2)$$

le spin total. Montrer que

$$S_{12} = 2\left[3Q^2 - \vec{\Sigma}^{\,2}\right] \qquad Q^2 = (\vec{\Sigma}\cdot\hat{r})^2$$

et que $Q^4 = Q^2$: Q^2 est un projecteur. En déduire que $S_{12}^2 = 4\vec{\Sigma}^2 - 2S_{12}$ et que les valeurs propres de S_{12} sont 0, 2 et -4.

9.7.17 Désintégration du Σ^0

La particule Σ^0, formée d'un quark up, d'un quark down et d'un quark étrange, de masse 1192 MeV/c^2 et de spin 1/2, se désintègre par une transition radiative en une particule Λ^0, également formée d'un quark up, d'un quark down et d'un quark étrange, de masse 1115 MeV/c^2 et de spin 1/2

$$\Sigma^0 \to \Lambda^0 + \gamma$$

Le Σ^0 est supposé au repos, son spin est quantifié suivant Oz et la projection de son spin sur cet axe est m ; l'impulsion $\vec{p}$ du photon est située dans le plan xOz et fait un angle θ avec l'axe Oz.

1. On suppose d'abord que le photon est émis dans la direction Oz ($\theta = 0$). Si m' est la projection du spin du Λ^0 sur Oz, montrer que les amplitudes non nulles sont (T est l'opérateur de transition)

$$\begin{aligned} a &= \langle D, m' = -\tfrac{1}{2}; \theta = 0|T|m = \tfrac{1}{2}\rangle \\ b &= \langle G, m' = \tfrac{1}{2}; \theta = 0|T|m = -\tfrac{1}{2}\rangle \end{aligned}$$

tandis que

$$\begin{aligned} c &= \langle D, m' = \tfrac{1}{2}; \theta = 0|T|m = \tfrac{1}{2}\rangle = 0 \\ d &= \langle G, m' = -\tfrac{1}{2}; \theta = 0|T|m = -\tfrac{1}{2}\rangle = 0 \end{aligned}$$

autrement dit $m' = m$ est interdit et les transitions permises correspondent à $m' = -m$ quand $\theta = 0$. La notation (D, G) spécifie l'état de polarisation circulaire droite (D) ou gauche (G) du photon.

2. L'opérateur de transition T est invariant dans l'opération parité. En déduire $|a| = |b|$. Si η est le produit des parités du Σ^0 et du Λ^0, aussi appelé parité relative des deux particules

$$\eta = \eta_{\Sigma^0}\eta_{\Lambda^0}$$

montrer que $a = \eta b$. L'expérience montre que $\eta = 1$, et donc $a = b$.

3. On suppose que la valeur initiale de la projection du spin du Σ^0 est $m = 1/2$. Soit $a_D^{m'}(\theta)$ et $a_G^{m'}(\theta)$ les amplitudes de transition où m' est la projection du spin du Λ^0 sur la direction de $\vec{p}$: c'est donc la valeur propre de $\vec{S}\cdot\hat{p}$. Calculer $a_D^{m'}$ et $a_G^{m'}$ en fonction de a et θ. Quelles sont les valeurs permises pour m' ?

9.7.18 Coefficients de Clebsch-Gordan du couplage $\vec{L}\cdot\vec{S}$

1. On considère la composition d'un moment angulaire orbital $\vec{L}$ et d'un spin 1/2 $\vec{S}$. Afin d'alléger les notations, on pose

$$|lsjm\rangle = \varphi(j,m) \qquad |lsm_lm_s\rangle = \psi(m_l,m_s)$$

et donc

$$\vec{J}^2\varphi(j,m) = j(j+1)\varphi(j,m) \qquad J_z\varphi(j,m) = m\varphi(j,m)$$

Le vecteur $\varphi(j,m)$ est une combinaison linéaire de $\psi(m-1/2,1/2)$ et de $\psi(m+1/2,-1/2)$. On examine d'abord le cas $j = l + 1/2$ en écrivant la combinaison linéaire

$$\varphi(l+1/2,m) = \alpha_m\psi(m-1/2,1/2)+\beta_m\psi(m+1/2,-1/2) \qquad |\alpha_m|^2+|\beta_m|^2 = 1$$

En appliquant $J_- = L_- + S_-$ aux deux membres de cette équation, montrer que

$$\begin{aligned}
J_-\varphi(l+1/2,m-1) &= \sqrt{(l+m+1/2)(l-m+3/2)}\ \varphi(l+1/2,m-1)\\
&= \sqrt{(l+m+1/2)(l-m+3/2)}\Big[\alpha_{m-1}\psi(m-3/2,1/2)\\
&\qquad +\beta_{m-1}\psi(m-1/2,-1/2)\Big]\\
&= \alpha_m\Big[\sqrt{(l+m-1/2)(l-m+3/2)}\ \psi(m-3/2,1/2)\\
&\qquad +\psi(m-1/2,-1/2)\Big]\\
&\qquad +\beta_m\sqrt{(l+m+1/2)(l-m+1/2)}\ \psi(m-1/2,-1/2)
\end{aligned}$$

En déduire
$$\frac{\alpha_m}{\sqrt{l+m+1/2}} = \frac{\alpha_{m-1}}{\sqrt{l+m-1/2}}$$
et en observant que ce rapport est indépendant de m
$$\alpha_m = \sqrt{\frac{l+m+1/2}{2l+1}}$$
En utilisant la condition $|\alpha_m|^2 + |\beta_m|^2 = 1$ et le fait que les coefficients de Clebsch-Gordan sont réels par convention, montrer que
$$\beta_m = \sqrt{\frac{l-m+1/2}{2l+1}}$$
Vérifier que le signe devant la racine carrée est bien $+$. On obtient donc pour les coefficients de Clebsch-Gordan.
$$C^{l,1/2}_{m-1/2,1/2;l+1/2,m} = \alpha_m = \sqrt{\frac{l+m+1/2}{2l+1}}$$
$$C^{l,1/2}_{m+1/2,-1/2;l+1/2,m} = \beta_m = \sqrt{\frac{l-m+1/2}{2l+1}}$$
En utilisant le fait que le vecteur $\varphi(l-1/2,m)$ est orthogonal au vecteur $\varphi(l+1/2,m)$, montrer que
$$\varphi(l-1/2,m) = \sqrt{\frac{l-m+1/2}{2l+1}}\,\psi(m-1/2) - \sqrt{\frac{l+m+1/2}{2l+1)}}\,\psi(m+1/2,-1/2)$$
et en déduire (par convention $C^{l,1/2}_{m+1/2,-1/2;l-1/2,m} > 0$)
$$C^{l,1/2}_{m-1/2,1/2;l-1/2,m} = -\beta_m = -\sqrt{\frac{l-m+1/2}{2l+1}}$$
$$C^{l,1/2}_{m+1/2,-1/2;l-1/2,m} = \alpha_m = \sqrt{\frac{l+m+1/2}{2l+1}}$$

9.7.19 Opérateurs tensoriels irréductibles

Un opérateur tensoriel irréductible d'ordre k, $T^{(k)}$ possède $(2k+1)$ composantes $T^{(k)}_q$
$$q = -k,\ -k+1, \ldots\ldots, k-1,\ k$$
et se transforme par une rotation $\mathcal{R}$ suivant
$$U(\mathcal{R})T^{(k)}_q U^\dagger(\mathcal{R}) = \sum_{q'} D^{(k)}_{q'q}(\mathcal{R})T^{(k)}_{q'}$$

Montrer que le vecteur

$$|kjqm\rangle = T_q^{(k)}|jm\rangle$$

se transforme par rotation comme un vecteur $|j_1 j_2 m_1 m_2\rangle$ avec $j_1 = k$, $j_2 = j$, $m_1 = q$, $m_2 = m$. En passant par l'intermédiaire des vecteurs

$$|kj\tilde{\jmath}\,\tilde{m}\rangle = \sum_{q+m=\tilde{m}} C^{kj}_{qm;\tilde{\jmath}\,\tilde{m}}|kjqm\rangle$$

déduire la forme générale du théorème de Wigner-Eckart

$$\langle j'm'|T_q^{(k)}|jm\rangle = C^{kj}_{qm;j'm'}\langle j'||T^{(k)}||jm\rangle$$

et donc

$$|j-k| \leq j' \leq j+k$$

9.8 Bibliographie

La présentation de ce chapitre, inspirée par celle de de Feynman *et al.* [1965], volume III, chapitres 17 et 18, met particulièrement l'accent sur les propriétés et l'utilisation des matrices de rotation. Pour une présentation plus classique, on se reportera à Messiah [1959], chapitre XIII, Cohen-Tannoudji *et al.* [1973], chapitre VII ou Basdevant et Dalibard [2001], chapitre 10. On trouvera de nombreuses applications à la physique des particules élémentaires dans le livre de Gasiorowicz [1966].

Chapitre 10

Oscillateur harmonique

L'oscillateur harmonique est un système d'une grande importance en mécanique classique, car il décrit les petites oscillations de systèmes physiques autour d'une position d'équilibre stable ; son importance n'est pas moindre en mécanique quantique. Pour fixer les idées, et afin de prendre l'exemple simple d'un mouvement à une dimension, examinons le cas des vibrations d'une molécule diatomique dont les deux noyaux ont des masses m_1 et m_2. On prend pour axe des x la droite joignant les deux noyaux et on note $x = x_1 - x_2$ la coordonnée de la particule relative (exercice 7.5.6). À l'équilibre, les deux noyaux se trouvent à une distance $x = x_0$. En physique classique, le hamiltonien de la particule relative s'écrit

$$H_{\rm cl} = \frac{p^2}{2m} + V(x) \tag{10.1}$$

où $m = m_1 m_2/(m_1+m_2)$ est la masse réduite. Développons $V(x)$ au voisinage de $x = x_0$:

$$V = V(x_0) + (x - x_0)V'(x_0) + \frac{1}{2}\,(x - x_0)^2 V''(x_0) + \cdots$$

La constante $V(x_0)$ est en général sans intérêt et on peut la prendre égale à zéro par une redéfinition du zéro d'énergie. Comme x_0 est une position d'équilibre, $V'(x_0) = 0$, et si cette position d'équilibre est stable, $V''(x_0) > 0$. Posant

$$q = x - x_0\,, \qquad C = V''(x_0)\,, \qquad \omega = \sqrt{\frac{C}{m}}$$

le hamiltonien (10.1) devient

$$\boxed{H_{\rm cl} = \frac{p^2}{2m} + \frac{1}{2}\,m\omega^2 q^2} \tag{10.2}$$

où ω est la fréquence des oscillations autour de la position d'équilibre.

Nous venons d'introduire l'exemple le plus simple, celui d'un oscillateur isolé, dont le traitement quantique fera l'objet de la section 10.1 dans une base particulière, celle des états propres de l'énergie. Une autre "base", celle des états cohérents, sera examinée dans la section suivante. Elle trouve de nombreuses applications en optique quantique. Il peut sembler surprenant que figure dans la dernière section de ce chapitre l'étude du mouvement d'une particule chargée dans un champ magnétique. Nous verrons que dans le cas d'un champ magnétique constant, les équations du mouvement se ramènent à celles de deux oscillateurs harmoniques. L'étude des niveaux d'énergie dans un champ magnétique, ou niveaux de Landau, sera précédée d'une définition de l'invariance de jauge locale, qui fixe la forme de l'interaction d'une particule chargée avec le champ électromagnétique.

10.1 L'oscillateur harmonique simple

10.1.1 Opérateurs de création et d'annihilation

Notre point de départ sera le hamiltonien (10.2). Ce hamiltonien se transpose en mécanique quantique si l'on interprète les quantités p et q comme des opérateurs : $p \to P$, $q \to Q$ et si l'on impose les relations de commutation canoniques

$$[Q, P] = \mathrm{i}\hbar I \tag{10.3}$$

Comme souvent en physique, il est utile de définir des quantités sans dimensions, et nous introduirons les opérateurs $\hat{P}$ et $\hat{Q}$ par

$$Q = \left(\frac{\hbar}{m\omega}\right)^{1/2} \hat{Q} \qquad P = (m\hbar\omega)^{1/2} \hat{P} \tag{10.4}$$

qui obéissent à la relation de commutation

$$[\hat{Q}, \hat{P}] = \mathrm{i}I \tag{10.5}$$

Nous allons construire les vecteurs propres de H par une méthode algébrique, suivant l'esprit de celle qui a été utilisée pour le moment angulaire. Le principe de cette méthode consiste à introduire des opérateurs a et $a^\dagger$, appelés respectivement *opérateur d'annihilation (ou de destruction)* et *opérateur de création* de l'oscillateur harmonique, qui feront passer d'une valeur propre de H à une autre, tout comme J_- et J_+ font passer d'une valeur propre de J_z à une autre. Nous définissons donc les opérateurs[1]

$$a = \frac{1}{\sqrt{2}} \left(\hat{Q} + \mathrm{i}\hat{P}\right) \tag{10.6}$$

$$a^\dagger = \frac{1}{\sqrt{2}} \left(\hat{Q} - \mathrm{i}\hat{P}\right) \tag{10.7}$$

1. Afin de nous conformer aux notations usuelles, nous dérogeons à notre règle selon laquelle les opérateurs sont notés par des lettres majuscules.

Un calcul immédiat permet d'établir les relations de commutation de a et $a^\dagger$

$$\boxed{[a, a^\dagger] = I} \tag{10.8}$$

ainsi que trois expressions utiles de H

$$\boxed{H = \frac{1}{2}\,\hbar\omega\left(\hat{P}^2 + \hat{Q}^2\right) = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) = \hbar\omega\left(N + \frac{1}{2}\right)} \tag{10.9}$$

Nous avons introduit l'opérateur N, ou opérateur *nombre de particules*[2]

$$\boxed{N = a^\dagger a} \tag{10.10}$$

qui vérifie les relations de commutation suivantes avec a et $a^\dagger$

$$[N, a] = -a \qquad\qquad [N, a^\dagger] = a^\dagger \tag{10.11}$$

Compte tenu de (10.9), il est équivalent de diagonaliser N ou H.

10.1.2 Diagonalisation du hamiltonien

Supposons que nous ayons trouvé un vecteur propre $|\nu\rangle$ de N, normalisable, mais pas nécessairement unitaire, de valeur propre ν

$$N|\nu\rangle = \nu|\nu\rangle$$

On doit avoir $\nu \geq 0$: en effet,

$$0 \leq ||a|\nu\rangle||^2 = \langle\nu|a^\dagger a|\nu\rangle = \langle\nu|N|\nu\rangle = \nu\langle\nu|\nu\rangle$$

ce qui implique que si $\nu = 0$, alors $a|\nu\rangle = 0$. Dans le cas contraire, $a|\nu\rangle$ est un vecteur de norme carrée $\nu\langle\nu|\nu\rangle$, et c'est un vecteur propre de N avec la valeur propre $(\nu - 1)$, car, en utilisant (10.11)

$$Na\big[|\nu\rangle\big] = a(N-1)|\nu\rangle = (\nu-1)\big[a|\nu\rangle\big]$$

Enfin, $a^\dagger|\nu\rangle$ est certainement un vecteur non nul, de norme carrée $(\nu+1)\langle\nu|\nu\rangle$, et c'est un vecteur propre de N avec la valeur propre $(\nu+1)$. En effet, d'une part

$$0 \leq ||a^\dagger|\nu||^2 = \langle\nu|aa^\dagger|\nu\rangle = \langle\nu|(N+1)|\nu\rangle = (\nu+1)\langle\nu|\nu\rangle$$

et d'autre part,

$$N\big[a^\dagger|\nu\rangle\big] = a^\dagger(N+1)|\nu\rangle = (\nu+1)\big[a^\dagger|\nu\rangle\big]$$

2. Cette terminologie sera justifiée pour les photons au § 17.1.1 et au chapitre 14 dans le cas général.

Si $\nu > 0$, nous avons vu que $a|\nu\rangle$ est vecteur propre de N avec la valeur propre $(\nu - 1)$. Si $(\nu - 1) = 0$, $a^2|\nu\rangle = 0$. Si $(\nu - 1) > 0$, on peut construire le vecteur non nul $a^2|\nu\rangle$, de valeur propre $(\nu - 2)$, et continuer le processus si $(\nu - 2) > 0$. La suite des vecteurs

$$a^0|\nu\rangle,\ a^1|\nu\rangle, a^2|\nu\rangle, \ldots, a^p|\nu\rangle \ldots$$

est une suite de vecteurs propres de N correspondant aux valeurs propres

$$\nu,\ \ \nu - 1,\ \ldots,\ (\nu - p)\ \ldots$$

Ceci montre que *ν est nécessairement un nombre entier*. Sinon, pour p suffisamment grand, $(\nu - p)$ deviendrait négatif et le vecteur $a^p|\nu\rangle$ serait de norme négative. Il est donc nécessaire que la série s'arrête pour une valeur *entière* $\nu = p$ telle que le vecteur $a^{p+1}|\nu\rangle = 0$.

La suite des vecteurs

$$(a^\dagger)^0|\nu\rangle,\ \ (a^\dagger)^1|\nu\rangle,\ \ (a^\dagger)^2|\nu\rangle,\ \ldots,\ (a^\dagger)^p|\nu\rangle \ldots$$

forme une suite de vecteurs propres de N correspondant aux valeurs propres

$$\nu,\ \ \nu + 1,\ \ldots, (\nu + p) \ldots$$

En résumé, *les valeurs propres de N sont des nombres entiers* :

$$n = \ 0,\ 1,\ 2,\ \ldots, n,\ \ldots$$

On note $|n\rangle$ les vecteurs propres de N correspondant à la valeur propre n

$$N|n\rangle = a^\dagger a|n\rangle = n|n\rangle \tag{10.12}$$

ou de façon équivalente pour H

$$H|n\rangle = \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle \tag{10.13}$$

Les valeurs propres E_n de l'énergie étiquetées par l'entier n sont de la forme

$$\boxed{E_n = \hbar\omega\left(n + \frac{1}{2}\right)} \tag{10.14}$$

Contrairement au cas de l'oscillateur classique, le niveau d'énergie E_0 du fondamental n'est pas zéro, ce qui correspondrait à une particule immobile à l'équilibre, mais $E_0 = \hbar\omega/2$: c'est *l'énergie de point zéro* de l'oscillateur harmonique. On peut en donner une explication qualitative grâce aux inégalités de Heisenberg (exercice 8.6.4). Il ne faut surtout pas confondre le vecteur propre $|0\rangle$ de l'état fondamental et le vecteur nul de l'espace de Hilbert $\mathcal{H}$, $|\varphi\rangle = 0$! On remarque aussi que les niveaux d'énergie sont équidistants : c'est

bien ce que l'on constate expérimentalement, en première approximation, pour les niveaux de vibration d'une molécule.

Les vecteurs $|n\rangle$ sont bien sûr orthogonaux si $n \neq n'$ et nous les supposons désormais unitaires. Il reste à montrer qu'ils ne sont pas dégénérés, qu'ils forment une base de l'espace de Hilbert $\mathcal{H}$, et avant tout que N a au moins un vecteur propre, ce qui n'est pas garanti pour un opérateur, même hermitien, dans un espace de dimension infinie ! Nous construirons explicitement dans la section suivante le vecteur $|0\rangle$ et nous montrerons qu'il est unique. Cela suffit à montrer que la série de vecteurs

$$|0\rangle, \ (a^\dagger)^1|0\rangle, \ (a^\dagger)^2|0\rangle, \ \ldots, (a^\dagger)^n|0\rangle \ldots \tag{10.15}$$

est unique. En effet, nous pouvons raisonner par récurrence en supposant le vecteur $|n\rangle$ non dégénéré. Soit $|n+1\rangle$, un vecteur propre de N correspondant à la valeur propre $(n+1)$: $N|n+1\rangle = (n+1)|n+1\rangle$. Alors, c étant un nombre complexe $\neq 0$,

$$a|n+1\rangle = c|n\rangle \Rightarrow a^\dagger a|n+1\rangle = ca^\dagger|n\rangle \Rightarrow |n+1\rangle = \frac{ca^\dagger}{n+1}\,|n\rangle$$

ce qui montre que $|n+1\rangle \propto a^\dagger|n\rangle$: si $|0\rangle$ est unique, le vecteur $|n\rangle$ est aussi unique à un facteur de phase près.

Comme dans le cas de la base standard $|jm\rangle$ du moment angulaire, il est commode de fixer un fois pour toutes la phase relative des vecteurs propres de H. Si $|n\rangle$ est de norme unité, le vecteur $a^\dagger|n\rangle$ est de norme $\sqrt{n+1}$ et par conséquent

$$a^\dagger|n\rangle = \mathrm{e}^{\,\mathrm{i}\alpha}\sqrt{n+1}\;|n+1\rangle$$

Le choix de phase le plus simple est $\alpha = 0$ et on a alors

$$a^\dagger|n\rangle = \sqrt{n+1}\;|n+1\rangle \tag{10.16}$$

$$a|n\rangle = \sqrt{n}\;|n-1\rangle \tag{10.17}$$

Les équations (10.16) et (10.17) mettent en évidence le rôle de création et de destruction de $a^\dagger$ et a : $a^\dagger$ augmente n d'une unité, a diminue n d'une unité. Les vecteurs $|n\rangle$ se déduisent de $|0\rangle$ par

$$\boxed{|n\rangle = \frac{1}{\sqrt{n!}}\,(a^\dagger)^n\,|0\rangle} \tag{10.18}$$

Il reste à montrer que les vecteurs $|n\rangle$ forment une base de $\mathcal{H}$: nous renvoyons ce point de rigueur à l'exercice 10.6.2.

10.1.3 Fonctions d'onde de l'oscillateur harmonique

En mécanique ondulatoire, le hamiltonien de l'oscillateur harmonique s'écrit

$$H = -\frac{\hbar^2}{2m}\,\frac{\mathrm{d}^2}{\mathrm{d}q^2} + \frac{1}{2}\,m\omega^2 q^2 \tag{10.19}$$

En définissant comme dans (10.4) la variable sans dimension u

$$q = \left(\frac{\hbar}{m\omega}\right)^{1/2} u \qquad -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}q} = -\mathrm{i}(\hbar m\omega)^{1/2}\,\frac{\mathrm{d}}{\mathrm{d}u} \tag{10.20}$$

le hamiltonien (10.19) devient

$$H = \frac{1}{2}\,\hbar\omega\left[-\frac{\mathrm{d}^2}{\mathrm{d}u^2} + u^2\right] \tag{10.21}$$

On aurait pu obtenir directement cette forme de H à partir de la première des équations (10.9) en remarquant que u et $-\mathrm{i}\mathrm{d}/\mathrm{d}u$ ne sont autres que les réalisations dans l'espace $L_u^{(2)}(\mathbb{R})$ des opérateurs $\hat{Q}$ et $\hat{P}$. Il serait possible de chercher directement les solutions de

$$H\varphi_n(u) = \frac{1}{2}\,\hbar\omega\left[-\frac{\mathrm{d}^2}{\mathrm{d}u^2} + u^2\right]\varphi_n(u) = E_n\varphi_n(u) \tag{10.22}$$

avec $\varphi_n(u) = \langle u|n\rangle$, mais nous nous limiterons à montrer l'unicité du vecteur $|0\rangle$ dont nous nous sommes servis ci-dessus. Comme $\langle u|0\rangle = \varphi_0(u)$, l'équation $a|0\rangle = 0$ devient

$$\langle u|a|0\rangle = \frac{1}{\sqrt{2}}\left[u + \frac{\mathrm{d}}{\mathrm{d}u}\right]\varphi_0(u) = 0$$

qui s'intègre immédiatement en

$$\varphi_0(u) = \frac{1}{\pi^{1/4}}\,\mathrm{e}^{-u^2/2} \tag{10.23}$$

Le facteur $\pi^{-1/4}$ assure que φ_0 est normalisé à l'unité. Cette solution est unique, ce qui prouve que les vecteurs propres donnés par la série (10.15) sont non dégénérés. Il est immédiat de vérifier que $\varphi_0(u)$ obéit à (10.22) avec la valeur propre $\hbar\omega/2$. La fonction $\varphi_0(u)$ vérifie bien la propriété caractéristique de la fonction d'onde d'un état fondamental : elle ne s'annule pas, ou de façon équivalente, elle n'a pas de nœuds.

Déterminons enfin la forme explicite des fonctions d'onde $\varphi_n(u) = \langle u|n\rangle$. Multiplions à gauche (10.18) écrit sous la forme

$$|n\rangle = \frac{1}{\sqrt{2^n\,n!}}\,\left(\hat{Q} - \mathrm{i}\hat{P}\right)^n|0\rangle$$

par le bra $\langle u|$

$$\varphi_n(u) = \langle u|n\rangle = \frac{1}{\pi^{1/4}}\,\frac{1}{\sqrt{2^n\,n!}}\,\left(u - \frac{\mathrm{d}}{\mathrm{d}u}\right)^n\,\mathrm{e}^{-u^2/2} \tag{10.24}$$

Les fonctions $\varphi_n(u)$ sont orthogonales pour $n \neq n'$ et normalisées à l'unité en raison de $\langle n|n'\rangle = \delta_{nn'}$. Les fonctions définies dans l'équation (10.24) sont reliées aux *polynômes de Hermite* $H_n(u)$

$$\mathrm{e}^{-u^2/2}\,H_n(u) = \left(u - \frac{\mathrm{d}}{\mathrm{d}u}\right)^n\,\mathrm{e}^{-u^2/2} \tag{10.25}$$

par

$$\varphi_n(u) = \frac{1}{\pi^{1/4}} \frac{1}{\sqrt{2^n\, n!}} \, \mathrm{e}^{-u^2/2} \, H_n(u) \tag{10.26}$$

Les premiers polynômes de Hermite sont

$$H_0(u) = 1 \qquad H_1(u) = 2u \qquad H_2(u) = 4u^2 - 2\ldots$$

Les polynômes de Hermite $H_n(u)$ sont des fonctions paires (impaires) de u pour n pair (impair), et il en est de même pour les $\varphi_n(u)$.

En résumé, on peut établir un "dictionnaire" permettant de passer de la "représentation n" du § 10.1.2 à la représentation du § 10.1.3 des états propres de H en termes de fonctions d'onde $\varphi_n(u)$. Dans le résumé ci-dessous, la première équation est écrite dans la base $|n\rangle$, et la seconde est l'équation équivalente en mécanique ondulatoire.

- Équation aux valeurs propres

$$\begin{aligned} \frac{1}{2}\left(\hat{P}^2 + \hat{Q}^2\right)|n\rangle &= \left(n + \frac{1}{2}\right)|n\rangle \Longleftrightarrow \\ \frac{1}{2}\left(-\frac{\mathrm{d}^2}{\mathrm{d}u^2} + u^2\right)\varphi_n(u) &= \left(n + \frac{1}{2}\right)\varphi_n(u) \end{aligned}$$

- Relations d'orthonormalisation

$$\langle n|m\rangle = \delta_{nm} \Longleftrightarrow \int_{-\infty}^{\infty} \mathrm{d}u\, \varphi_n^*(u)\varphi_m(u) = \delta_{nm}$$

- Relation de fermeture

$$\sum_n |n\rangle\langle n| = I \Longleftrightarrow \sum_n \varphi_n(u)\varphi_n^*(v) = \delta(u - v)$$

La conjugaison complexe est en fait superflue car les fonctions $\varphi_n(u)$ sont réelles.

10.2 États cohérents

10.2.1 Définition et propriétés élémentaires

Les *états cohérents*, ou *états semi-classiques*, sont des états quantiques remarquables de l'oscillateur harmonique : dans ces états, les valeurs moyennes des opérateurs position et impulsion ont des propriétés aussi proches que possibles des valeurs classiques de la position $q(t)$ et de l'impulsion $p(t)$. L'exercice 10.4.3 montre que l'expression des états cohérents découle de cette propriété. Nous les définirons ci-dessous *a priori* et nous vérifierons qu'ils ont

bien les propriétés annoncées. Soit $z(t)$, un nombre complexe, combinaison de $q(t)$ et $p(t)$

$$z(t) = \sqrt{\frac{m\omega}{2\hbar}}\, q(t) + \frac{\mathrm{i}}{\sqrt{2m\hbar\omega}}\, p(t) = \frac{1}{\sqrt{2}}\,(\hat{q}(t) + \hat{p}(t)) \tag{10.27}$$

si l'on relie (q, p) à $(\hat{q}, \hat{p})$ de la même manière que (Q, P) à $(\hat{Q}, \hat{P})$ (10.4) : $z(t)$ est "l'équivalent classique" de $a(t)$. À partir des équations du mouvement classiques

$$\frac{\mathrm{d}q(t)}{\mathrm{d}t} = \frac{1}{m}\, p(t) \qquad \frac{\mathrm{d}p(t)}{\mathrm{d}t} = -m\omega^2 q(t) \tag{10.28}$$

on montre que $z(t)$ vérifie l'équation différentielle

$$\frac{\mathrm{d}z}{\mathrm{d}t} = -\mathrm{i}\omega z(t) \tag{10.29}$$

qui a pour solution

$$z(t) = z_0 \mathrm{e}^{-\mathrm{i}\omega t}$$

Le nombre complexe $z(t)$ décrit une trajectoire circulaire dans le plan complexe en z avec une vitesse angulaire uniforme. On déduit de $z(t)$ la position $q(t)$, l'impulsion $p(t)$ et l'énergie de l'oscillateur

$$\begin{aligned} q(t) &= \sqrt{\frac{2\hbar}{m\omega}}\, \mathrm{Re}\, z(t) \\ p(t) &= \sqrt{2m\hbar\omega}\, \mathrm{Im}\, z(t) \\ E &= \hbar\omega |z_0|^2 \end{aligned} \tag{10.30}$$

Il est facile de montrer que la valeur moyenne $\langle a \rangle(t)$ de l'opérateur d'annihilation a obéit à la même équation différentielle que $z(t)$ (exercice 10.4.3). Ceci suggère de chercher les vecteurs propres de l'opérateur a, dont nous allons montrer qu'ils existent[3], car les valeurs propres correspondantes obéiront alors à (10.29). Ces vecteurs propres sont précisément les états cohérents. Un *état cohérent* $|z\rangle$ est défini par

$$\boxed{|z\rangle = \mathrm{e}^{-|z^2|/2} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}}\, |n\rangle = \mathrm{e}^{-|z|^2/2}\, \mathrm{e}^{za^\dagger}\, |0\rangle} \tag{10.31}$$

Énonçons quelques propriétés des états cohérents et vérifions tout d'abord que $|z\rangle$ est vecteur propre de a.

- L'état cohérent $|z\rangle$ est un vecteur propre de l'opérateur d'annihilation (non hermitien) a avec la valeur propre z

$$\boxed{a|z\rangle = z|z\rangle} \tag{10.32}$$

3. Il n'est pas évident *a priori* que a, qui n'est pas un opérateur hermitien, ait des vecteurs propres, et encore moins que ces vecteurs propres forment une base de $\mathcal{H}$.

Pour le montrer, on peut utiliser directement (10.31), mais on peut aussi se servir de l'identité (2.70) de l'exercice 2.5.11 qui s'écrit ici

$$\begin{aligned} \mathrm{e}^{a^\dagger z} a\, \mathrm{e}^{-a^\dagger z} &= a + z[a^\dagger, a] = a - z \\ \mathrm{e}^{a^\dagger z} a &= (a-z)\mathrm{e}^{a^\dagger z} \end{aligned}$$

Il suffit d'appliquer les deux membres de cette dernière équation sur le vecteur $|0\rangle$ pour obtenir (10.32).

- Le vecteur $|z\rangle$ est unitaire : $\langle z|z\rangle = 1$ et le module carré du produit scalaire $\langle z|z'\rangle$

$$|\langle z|z'\rangle|^2 = \exp\left(-|z-z'|^2\right) \tag{10.33}$$

est une mesure de la “distance” entre les deux états cohérents.

- La distribution des valeurs de n suit une loi de Poisson

$$\mathsf{p}(n) = |\langle n|z\rangle|^2 = \frac{|z|^{2n}}{n!}\mathrm{e}^{-|z|^2} \tag{10.34}$$

ce qui donne pour la valeur moyenne $\langle n\rangle = |z|^2$ et pour la dispersion $\Delta n = |z|$.

- L'action de $\exp(\lambda N)$ sur un état cohérent, où λ est un nombre imaginaire pur ($|\exp\lambda| = 1$), redonne un état cohérent

$$\begin{aligned} \mathrm{e}^{\lambda N}|z\rangle &= \mathrm{e}^{\lambda N}\mathrm{e}^{-|z|^2/2}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,|n\rangle = \mathrm{e}^{-|z|^2/2}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,\mathrm{e}^{\lambda n}|n\rangle \\ &= \mathrm{e}^{-|z|^2/2}\sum_{n=0}^{\infty}\frac{(\mathrm{e}^{\lambda}z)^n}{\sqrt{n!}}\,|n\rangle = |\mathrm{e}^{\lambda}z\rangle \end{aligned} \tag{10.35}$$

La relation $|\exp\lambda| = 1$ a été utilisée uniquement pour obtenir la dernière égalité.

- Les états cohérents forment une base “surcomplète”

$$\int \frac{\mathrm{dRe}\, z\, \mathrm{dIm}\, z}{\pi}\,|z\rangle\langle z| = I \tag{10.36}$$

Pour montrer cette identité, on la prend en sandwich entre le bra $\langle n|$ et le ket $|m\rangle$; posant $z = \rho\exp(\mathrm{i}\theta)$

$$\begin{aligned} \int \frac{\mathrm{dRe}\, z\, \mathrm{dIm}\, z}{\pi}\,\langle n|z\rangle\langle z|m\rangle &= \int_0^\infty \rho\,\mathrm{d}\rho\int_0^{2\pi}\frac{\mathrm{d}\theta}{\pi}\,\frac{z^n z^{*m}}{\sqrt{n!m!}}\,\mathrm{e}^{-\rho^2} \\ &= \int_0^\infty \rho\,\mathrm{d}\rho\int_0^{2\pi}\frac{\mathrm{d}\theta}{\pi}\,\frac{\rho^{n+m}}{\sqrt{n!m!}}\,\mathrm{e}^{\mathrm{i}(n-m)\theta}\mathrm{e}^{-\rho^2} \\ &= \delta_{nm} \end{aligned}$$

où nous avons utilisé le changement de variables $\rho^2 = u$ et

$$\int_0^\infty \mathrm{d}u\, u^n \mathrm{e}^{-u} = n!$$

Une conséquence directe de (10.36) est que les “éléments de matrice diagonaux” $\langle z|A|z\rangle$ suffisent à définir complètement un opérateur A (exercice 10.4.3).

Ces propriétés permettent de calculer aisément les valeurs moyennes

$$\begin{aligned}
\langle z|Q|z\rangle &= \sqrt{\frac{\hbar}{2m\omega}}\,\langle z|\left(a+a^\dagger\right)|z\rangle = \sqrt{\frac{2\hbar}{m\omega}}\,\mathrm{Re}\,z \\
\langle z|P|z\rangle &= \sqrt{2\hbar m\omega}\,\mathrm{Im}\,z \\
\langle z|H|z\rangle &= \hbar\omega\left(|z|^2+\frac{1}{2}\right)
\end{aligned} \tag{10.37}$$

Ceci est le résultat classique (10.30), si l'on ignore l'énergie de point zéro $\hbar\omega/2$ dans l'expression de $\langle H\rangle$. La valeur moyenne de N est $\langle z|N|z\rangle = |z|^2$, résultat que l'on prouve directement à partir de (10.10) ou de la distribution de Poisson (10.34). De plus, si l'état de l'oscillateur harmonique est un état cohérent au temps $t=0$, cette propriété est conservée par l'évolution temporelle. Supposons en effet que l'oscillateur se trouve au temps $t=0$ dans l'état cohérent $|\varphi(t=0)\rangle = |z_0\rangle$ et calculons $|\varphi(t)\rangle$

$$|\varphi(t)\rangle = \mathrm{e}^{-\mathrm{i}Ht/\hbar}\,|z_0\rangle = \mathrm{e}^{-\mathrm{i}\omega t/2}\,\mathrm{e}^{-\mathrm{i}\omega Nt}\,|z_0\rangle = \mathrm{e}^{-\mathrm{i}\omega t/2}\,|z_0\mathrm{e}^{-\mathrm{i}\omega t}\rangle \tag{10.38}$$

où nous avons utilisé (11.35). Au facteur de phase $\exp(-\mathrm{i}\omega t/2)$ près, on retrouve l'évolution classique $z(t) = z_0\exp(-\mathrm{i}\omega t)$. Si l'on part d'un état cohérent au temps $t=0$, l'évolution des valeurs moyennes $\langle Q\rangle$, $\langle P\rangle$ et $\langle H\rangle$ suit très exactement l'évolution classique de $q(t)$, $p(t)$ et E. Nous avons donc montré que les valeurs moyennes dans un état cohérent suivent les lois classiques.

Il est également instructif de calculer les dispersions. Évaluons par exemple $\langle Q^2\rangle$ dans l'état cohérent $|z\rangle$

$$\begin{aligned}
\langle Q^2\rangle_z &= \frac{\hbar}{2m\omega}\,\langle z|a^2+(a^\dagger)^2+aa^\dagger+a^\dagger a|z\rangle \\
&= \frac{\hbar}{2m\omega}\,\langle z|a^2+(a^\dagger)^2+2a^\dagger a+1|z\rangle \\
&= \frac{\hbar}{2m\omega}\,\left[1+(z+z^*)^2\right] = \frac{\hbar}{2m\omega}\,\left[1+4(\mathrm{Re}\,z)^2\right]
\end{aligned}$$

Un calcul analogue (exercice 10.4.3) donne $\langle P^2\rangle$ et $\langle H^2\rangle$, d'où l'on déduit les dispersions[4] dans l'état cohérent $|z\rangle$

$$\boxed{\Delta_z Q = \sqrt{\frac{\hbar}{2m\omega}} \qquad \Delta_z P = \sqrt{\frac{m\hbar\omega}{2}} \qquad \Delta_z H = \hbar\omega|z|} \tag{10.39}$$

La dispersion $\Delta_z H$ peut être obtenue en utilisant $\Delta H = \hbar\omega\Delta_z N$ et $\Delta_z N = \Delta n = |z|$ d'après (10.34) : on peut aussi calculer directement $\langle z|N^2|z\rangle$. On note

4. Nous utilisons indifféremment pour les dispersions les notations $(\Delta P, \Delta Q)$ ou $(\Delta p, \Delta q)$ car aucune ambiguïté n'est possible.

que *l'inégalité de Heisenberg est saturée dans un état cohérent* : $\Delta_z Q\,\Delta_z P = \hbar/2$ et que pour $|z| \gg 1$

$$\frac{\Delta_z H}{\langle H\rangle} \simeq \frac{1}{|z|} \to 0 \quad \text{si} \quad |z| \to \infty$$

En résumé, pour $|z| \gg 1$, les dispersions autour des valeurs moyennes sont les plus faibles possibles.

10.2.2 Opérateurs de déplacement et de phase

Au lieu de construire les états cohérents à partir de (10.31), il est souvent commode d'utiliser l'*opérateur de déplacement* $D(z)$ défini par

$$D(z) = \exp(-z^* a + z a^\dagger) = \exp\left(-\frac{1}{2}\,|z|^2\right)\exp(z a^\dagger)\,\exp(-z^* a) \tag{10.40}$$

où nous avons utilisé (2.71) pour obtenir la seconde expression dans (10.40). Cette seconde expression montre que l'action de $D(z)$ sur l'état $|0\rangle$ donne bien l'état cohérent $|z\rangle$

$$D(z)|0\rangle = \exp\left(-\frac{1}{2}\,|z|^2\right)\exp(z a^\dagger)\,|0\rangle = |z\rangle \tag{10.41}$$

L'opérateur conjugué hermitien de $D(z)$ est $D(-z)$

$$D^\dagger(z) = \exp\left(-\frac{1}{2}\,|z|^2\right)\exp(-z a^\dagger)\,\exp(z^* a) = D(-z)$$

ce qui montre que $D(z)$ est unitaire : $D^\dagger(z)D(z) = I$.

La propriété la plus utile de l'opérateur de déplacement est la suivante : le produit de deux opérateurs $D(z_1)$ et $D(z_2)$ est $D(z_1 + z_2)$ à un facteur de phase près. Pour le montrer, définissons les opérateurs A_1 et A_2 par

$$A_1 = z_1 a^\dagger - z_1^* a \qquad\qquad A_2 = z_2 a^\dagger - z_2^* a$$

et calculons leur commutateur

$$[A_1, A_2] = z_1 z_2^* - z_1^* z_2 = 2\,\mathrm{i}\,\mathrm{Im}(z_1 z_2^*)$$

Utilisons cette fois (2.72)

$$\begin{aligned} D(z_1)D(z_2) &= \exp(A_1)\exp(A_2) \\ &= \exp[\mathrm{i}\,\mathrm{Im}(z_1 z_2^*)\exp\left[(z_1 + z_2)\,a^\dagger - (z_1^* + z_2^*)\,a\right] \\ &= \exp[\mathrm{i}\,\mathrm{Im}(z_1 z_2^*)]D(z_1 + z_2) \end{aligned} \tag{10.42}$$

Cette propriété permet de montrer simplement le résultat important suivant : si un oscillateur harmonique initialement dans son état fondamental est couplé

pendant un intervalle de temps $[t_1, t_2]$ à une force classique $F(t)$, il se trouvera pour $t > t_2$ dans un état cohérent déterminé par la transformée de Fourier $\tilde{F}(\omega)$ de $F(t)$ (exercice 10.4.4)

$$\tilde{F}(\omega) = \int_{t_1}^{t_2} \mathrm{e}^{\mathrm{i}\omega t}\, F(t)\mathrm{d}t \tag{10.43}$$

Nous allons maintenant préciser la correspondance entre le mouvement classique et la description quantique par un état cohérent. Nous avons déjà vu que la trajectoire classique dans le plan $\{q(t), p(t)\}$ était déterminée par la valeur initiale z_0 de z (*cf.* (10.27)) ; z_0 est un nombre complexe $|z_0|\exp(\mathrm{i}\theta_0)$, où le module $|z_0|$ fixe l'énergie de l'oscillateur dans le cas classique. Dans le cas quantique d'un oscillateur dans un état cohérent $|z_0\rangle$, $|z_0|$ fixe la valeur moyenne de l'énergie. Il serait également intéressant de pouvoir donner une interprétation de la phase $\phi(t) = \theta_0 - \omega t$. Pour ce faire, examinons le portrait de phase de l'oscillateur, c'est-à-dire la trajectoire des valeurs moyennes dans le plan $\{q, p\}$. Dans le cas d'un état cohérent, ce portrait de phase est donné par les valeurs moyennes $\langle Q\rangle(t)$ et $\langle P\rangle(t)$ dans l'état cohérent $|z(t)\rangle$, avec $z(t) = z_0\exp(-\mathrm{i}\omega t)$, et d'après (10.37) en prenant $\theta_0 = 0$ pour simplifier les formules

$$\begin{aligned} \langle z(t)|Q|z(t)\rangle &= \sqrt{\frac{2\hbar}{m\omega}}\,|z_0|\cos\omega t \\ \langle z(t)|P|z(t)\rangle &= -\sqrt{2\hbar m\omega}\,|z_0|\sin\omega t \end{aligned} \tag{10.44}$$

Le point $\{\langle Q\rangle(t), \langle P\rangle(t)\}$ décrit donc une ellipse. Pour se ramener à une trajectoire circulaire, il suffit d'utiliser les opérateurs sans dimension $\hat{Q}$ et $\hat{P}$ (10.4)

$$\begin{aligned} \langle\hat{Q}\rangle(t) &= \sqrt{2}\,|z_0|\cos\omega t \\ \langle P\rangle(t) &= -\sqrt{2}\,|z_0|\sin\omega t \end{aligned} \tag{10.45}$$

Le point $\{\langle\hat{Q}\rangle(t), \langle\hat{P}\rangle(t)\}$ décrit un cercle de rayon $\sqrt{2}\,|z_0|$ dans le plan $\{\hat{q}, \hat{p}\}$. D'après (10.39), les dispersions $\Delta_z\hat{Q}$ et $\Delta_z\hat{P}$ sont indépendantes de z

$$\Delta_z\hat{Q} = \Delta_z\hat{P} = \frac{1}{\sqrt{2}}$$

On observe qu'au temps $t = 0$, le point représentatif du mouvement se trouve en $\{\hat{q} = \sqrt{2}\,|z_0|, \hat{p} = 0\}$, avec une dispersion horizontale et verticale de l'ordre de $1/\sqrt{2}$. Au temps $t = \pi/2\omega$, il se trouve en $\{\hat{q} = 0, \hat{p} = -\sqrt{2}\,|z_0|\}$ avec la même dispersion (figure 10.1). Le point représentatif fluctue donc autour de la position moyenne définie par (10.45). Ceci correspond à des fluctuations du rayon de la trajectoire et de sa phase : le rayon fluctue autour de sa valeur moyenne $|z_0|\sqrt{2}$ et la phase autour de sa valeur moyenne $\langle\phi\rangle = -\omega t$. Cette fluctuation de phase peut s'écrire

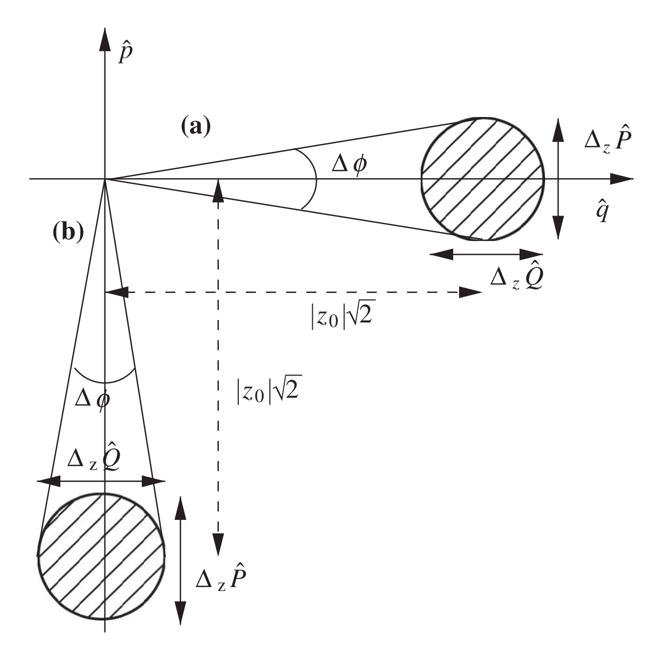


FIG. 10.1 – Représentation d'un état cohérent à $t = 0$ (a) et $t = \pi/2\omega$ (b).

$$\Delta_z\phi = \frac{1}{\sqrt{2}}\,\frac{1}{|z_0|\sqrt{2}} = \frac{1}{2|z_0|}$$

Par ailleurs, la loi de Poisson (10.34) relie $|z_0|$ et la fluctuation $\Delta_z N$ de N dans l'état cohérent : comme nous l'avons montré, $\Delta_z N = |z_0|$, de sorte que le produit des dispersions $\Delta_z\phi$ et $\Delta_z N$ dans un état cohérent $|z\rangle$ obéit à

$$\boxed{\Delta_z\phi\,\Delta_z N \simeq \frac{1}{2}} \tag{10.46}$$

Cette relation rappelle bien sûr une inégalité de Heisenberg, et l'analogie avec la relation de commutation canonique $[\hat{Q}, \hat{P}] = \mathrm{i}I$ suggère qu'il pourrait exister un opérateur de phase Φ canoniquement conjugué de N et obéissant à $[\Phi, N] = \mathrm{i}I$. En fait, il est impossible de construire un tel opérateur de phase qui soit auto-adjoint, en raison la périodicité de la variable ϕ (*cf.* l'exercice 6.4.3). On montre dans l'exercice 10.4.5 que l'opérateur d'annihilation possède une quasi-représentation en module et phase, $a = A\exp(\mathrm{i}\Phi)$; cependant, si $\exp(\mathrm{i}\Phi)$ (non unitaire !) existe, Φ lui-même n'existe pas comme opérateur hermitien (plus exactement auto-adjoint). L'opérateur $\exp(\mathrm{i}\Phi)$ sera noté E, avec

$$E = \sum_{n=0}^{\infty} |n\rangle\langle n+1| \qquad E^\dagger = \sum_{n=0}^{\infty} |n+1\rangle\langle n| \tag{10.47}$$

Les opérateurs E et $E^\dagger$ vérifient

$$EE^\dagger = I \qquad E^\dagger E = I - |0\rangle\langle 0| \tag{10.48}$$

Les *états de phase* $|\phi\rangle$ définis par

$$|\phi\rangle = \sum_{n=0}^{\infty} \mathrm{e}^{\,in\phi}\,|n\rangle \tag{10.49}$$

sont vecteurs propres de E avec la valeur propre $\exp(\mathrm{i}\phi)$

$$E|\phi\rangle = \mathrm{e}^{\,\mathrm{i}\phi}\,|\phi\rangle$$

Comme l'opérateur E n'est pas hermitien, les états $|\phi\rangle$ ne sont pas orthogonaux, mais ils vérifient la relation de fermeture

$$\int \frac{\mathrm{d}\phi}{2\pi}\,|\phi\rangle\langle\phi| = I$$

Compte tenu de la représentation $a = AE$ et de $a|z\rangle = z|z\rangle$, on s'attend à ce que la distribution $\mathsf{p}(\phi)$ de ϕ soit centrée autour de θ pour un état cohérent $|z\rangle$, $z = |z|\exp(\mathrm{i}\theta)$. Cherchons cette distribution $\mathsf{p}(\phi)$,

$$\mathsf{p}(\phi) = \frac{1}{2\pi}\,|\langle\phi|z\rangle|^2$$

où le facteur (2π) assure que $\mathsf{p}(\phi)$ est normalisé à l'unité pour une distribution uniforme dans l'intervalle $[0, 2\pi]$. L'amplitude $\langle\phi|z\rangle$ est donnée par

$$\langle\phi|z\rangle = \mathrm{e}^{-|z|^2/2}\sum_{n=0}^{\infty}\mathrm{e}^{-\mathrm{i}n(\phi-\theta)}\,\frac{|z|^{n/2}}{\sqrt{n!}}$$

Pour $|z| \gg 1$, on peut utiliser la formule de Stirling

$$n! \simeq \left(\frac{n}{\mathrm{e}}\right)^n\sqrt{2\pi n}$$

et écrire

$$\mathrm{e}^{-|z|^2/2}\,\frac{|z|^{n/2}}{\sqrt{n!}} \simeq \left(\frac{1}{2\pi|z|^2}\right)^{1/4}\exp\left[-\frac{1}{4|z|^2}(n-|z|^2)^2\right]$$

On peut en outre remplacer la somme sur n par une intégrale. L'amplitude $\langle\phi|z\rangle$ est alors donnée par la tranformée de Fourier d'une gaussienne

$$\langle\phi|z\rangle \simeq \left(8\pi|z|^2\right)^{1/4}\,\mathrm{e}^{-|z|^2(\phi-\theta)^2}$$

et $\mathsf{p}(\phi)$ s'en déduit immédiatement

$$\mathsf{p}(\phi) \simeq \sqrt{\frac{2|z|^2}{\pi}}\,\mathrm{e}^{-2|z|^2(\phi-\theta)^2} \tag{10.50}$$

On vérifie que $\int \mathsf{p}(\phi)\mathrm{d}\phi = 1$. L'équation (10.50) montre que la phase ϕ est bien centrée autour de θ avec une dispersion $\Delta_z\phi = 1/2|z|$, en accord avec (10.46).

10.3 Mouvement dans un champ magnétique

10.3.1 Invariance de jauge locale

Une application importante de l'oscillateur harmonique concerne le mouvement d'une particule de charge q dans un champ magnétique constant. Établissons en premier lieu la forme de l'interaction entre une particule chargée et un champ électromagnétique *classique* $(\vec{E}, \vec{B})$. En électrodynamique classique, la densité de charge électrique $\rho_{\text{em}}(\vec{r}, t)$ et la densité de courant

$$\vec{\jmath}_{\text{em}}(\vec{r}, t) = \rho_{\text{em}}(\vec{r}, t)\vec{v}(\vec{r}, t) \tag{10.51}$$

vérifient l'équation de continuité

$$\frac{\partial \rho_{\text{em}}}{\partial t} + \vec{\nabla} \cdot \vec{\jmath}_{\text{em}} = 0 \tag{10.52}$$

Nous souhaitons généraliser l'expression de ce courant à la physique quantique. Nous avons établi au chapitre 8 l'expression du courant de particules (8.123)

$$\begin{aligned}
\vec{\jmath}(\vec{r}, t) &= \text{Re}\left(\varphi^*(\vec{r}, t)\left[\frac{-\mathrm{i}\hbar}{m}\vec{\nabla}\right]\varphi(\vec{r}, t)\right) \\
&= \left(\varphi^*(\vec{r}, t)\left[\frac{-\mathrm{i}\hbar}{2m}\vec{\nabla}\right]\varphi(\vec{r}, t) - \varphi(\vec{r}, t)\left[\frac{-\mathrm{i}\hbar}{2m}\vec{\nabla}\right]\varphi^*(\vec{r}, t)\right)
\end{aligned} \tag{10.53}$$

Le courant électromagnétique créé par le mouvement d'une particule quantique chargée de charge q devrait être *a priori* $j_{\text{em}} = q\vec{\jmath}$, la densité de charge ρ_{em} étant $q|\varphi|^2$. Cette forme du courant de particules obéit à l'équation de continuité lorsque la fonction d'onde $\varphi(\vec{r}, t)$ vérifie l'équation de Schrödinger

$$\mathrm{i}\hbar\frac{\partial \varphi}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\varphi$$

et le courant électromagnétique associé

$$\rho_{\text{em}} = q|\varphi|^2, \quad \vec{\jmath}_{\text{em}} = q\vec{\jmath}$$

obéit à (10.52). Cependant, nous allons voir que l'expression (10.53) du courant doit être modifiée en présence d'un potentiel vecteur. L'expression (10.53) du courant est invariante dans une *transformation de jauge globale*, qui consiste à multiplier φ par un facteur de phase

$$\boxed{\varphi(\vec{r}, t) \to \varphi'(\vec{r}, t) = \exp\left(-\mathrm{i}\frac{q}{\hbar}\Lambda\right)\varphi(\vec{r}, t) = \Omega\varphi(\vec{r}, t)} \tag{10.54}$$

où Λ est un nombre réel. Lorsque Λ est une fonction de $\vec{r}$ et de t, on a affaire à une *transformation de jauge locale*. Nous allons déduire la forme

du courant d'un principe d'invariance de jauge locale, qui pourra sembler *a priori* arbitraire. En fait, ce principe est très général, car on pense aujourd'hui que toutes les interactions fondamentales de la physique se déduisent d'un tel principe (exercice 10.4.7). Une transformation de jauge locale s'obtient en remplaçant la constante Λ dans (10.54) par une *fonction* de $\vec{r}$ et de t

$$\boxed{\varphi(\vec{r},t) \to \varphi'(\vec{r},t) = \exp\left(-\mathrm{i}\frac{q}{\hbar}\,\Lambda(\vec{r},t)\right)\varphi(\vec{r},t) = \Omega(\vec{r},t)\varphi(\vec{r},t)} \tag{10.55}$$

Cette transformation est manifestement unitaire. Il est immédiat de vérifier que le courant (10.53) n'est pas invariant dans une transformation de jauge locale, car le gradient agit sur $\exp(\mathrm{i}q\Lambda/\hbar)$. Nous allons modifier l'expression du courant en substituant au gradient $\vec{\nabla}$ la *dérivée covariante* $\vec{D}$

$$\boxed{-\mathrm{i}\hbar\vec{D} = -\mathrm{i}\hbar\vec{\nabla} - q\vec{A}} \tag{10.56}$$

Contrairement à la dérivée ordinaire, la dérivée covariante a une action simple lorsque l'on effectue une transformation de jauge locale (10.55)

$$\begin{aligned}
-\mathrm{i}\hbar\vec{D}\varphi = -\mathrm{i}\hbar\vec{D}(\Omega^{-1}\varphi') &= (-\mathrm{i}\hbar\vec{\nabla} - q\vec{A})\exp\left(\mathrm{i}\frac{q}{\hbar}\,\Lambda(\vec{r},t)\right)\varphi'(\vec{r},t) \\
&= \Omega^{-1}(-\mathrm{i}\hbar\vec{\nabla} - q\vec{A} + q\vec{\nabla}\Lambda)\varphi' \\
&= \Omega^{-1}(-\mathrm{i}\hbar\vec{\nabla} - q\vec{A}')\varphi' = \Omega^{-1}(-\mathrm{i}\hbar\vec{D}'\varphi')
\end{aligned} \tag{10.57}$$

où $\vec{D}'$ est la dérivée covariante calculée avec le potentiel vecteur transformé

$$\vec{A}' = \vec{A} - \vec{\nabla}\Lambda \tag{10.58}$$

Les dérivées covariantes $\vec{D}$ et $\vec{D}'$ sont physiquement équivalentes car $\vec{A}$ est $\vec{A}'$ le sont. L'expression du courant devient invariante par transformation de jauge locale si, au lieu de la dérivée ordinaire, on se sert de la dérivée covariante

$$\boxed{\vec{\jmath}(\vec{r},t) = \mathrm{Re}\left(\varphi^*(\vec{r},t)\left[\frac{-\mathrm{i}\hbar}{m}\vec{\nabla} - \frac{q}{m}\vec{A}\right]\varphi(\vec{r},t)\right) = \mathrm{Re}\left(\varphi^*(\vec{r},t)\left[\frac{-\mathrm{i}\hbar}{m}\vec{D}\varphi\right]\right)} \tag{10.59}$$

En effet, si l'on exprime φ en fonction de φ' en utilisant (10.55) et (10.59)

$$\begin{aligned}
\vec{\jmath}(\vec{r},t) &= \mathrm{Re}\left(\varphi'^*(\vec{r},t)\Omega\,\Omega^{-1}\left[\frac{-\mathrm{i}\hbar}{m}\vec{D}'\varphi'\right]\right) \\
&= \mathrm{Re}\left(\varphi'^*(\vec{r},t)\left[\frac{-\mathrm{i}\hbar}{m}\vec{D}'\varphi'\right]\right) = \vec{\jmath}'(\vec{r},t)
\end{aligned}$$

Ceci suggère que l'opérateur vitesse $\mathrm{d}\vec{R}/\mathrm{d}t$ n'est pas simplement $\mathrm{d}\vec{R}/\mathrm{d}t = \vec{P}/m = -(\mathrm{i}\hbar/m)\vec{\nabla}$ mais plutôt

$$\frac{\mathrm{d}\vec{R}}{\mathrm{d}t} = -\frac{\mathrm{i}\hbar}{m}\,\vec{D} = -\frac{\mathrm{i}\hbar}{m}\,\vec{\nabla} - \frac{q}{m}\,\vec{A} \tag{10.60}$$

Sachant que l'opérateur vitesse est donné par le commutateur de $\vec{R}$ avec le hamiltonien, examinons sa composante x. D'après (7.61) et l'expression (10.60) de $\mathrm{d}\vec{R}/\mathrm{d}t$

$$\dot{X} = \frac{\mathrm{i}}{\hbar}\,[H, X] = \frac{1}{m}\,(P_x - qA_x)$$

ce qui donne, en reprenant le raisonnement de la section 7.4, la forme la plus générale de H

$$H = \frac{1}{2m}\left(\vec{P} - q\vec{A}\right)^2 + q\overline{V} = \frac{1}{2m}\left(-\mathrm{i}\hbar\vec{\nabla} - q\vec{A}\right)^2 + q\overline{V} = \frac{1}{2m}(-\mathrm{i}\hbar\vec{D})^2 + q\overline{V} \tag{10.61}$$

où $V = q\overline{V}$ est une fonction arbitraire de $\vec{R}$ et t. Exiger l'invariance de jauge locale du courant permet de retrouver la forme générique (7.73) du hamiltonien compatible avec l'invariance galiléenne. La substitution $-\mathrm{i}\hbar\vec{\nabla} \to -\mathrm{i}\hbar\vec{D}$ dans l'équation de Schrödinger, écrite dans un premier temps en l'absence de champ électromagnétique, et qui permet d'écrire cette équation en présence d'un tel champ est appelée *couplage minimal*[5]. La prescription du couplage minimal s'étend aux théories de jauge non abéliennes (exercice 10.4.7) et permet d'écrire toutes les interactions du modèle standard de la physique des particules élémentaires entre les particules de spin 1/2 (“particules de matière”) et celles de spin 1 (bosons de jauge) définies au § 1.1.3.

En mécanique analytique (exercice 7.5.8), on montre que le hamiltonien dont se déduit la force de Lorentz (1.11) est

$$H_{\mathrm{cl}} = \frac{1}{2m}\left(\vec{p} - q\vec{A}\right)^2 + q\overline{V}$$

Un autre façon d'obtenir (10.61) est de partir de cette forme classique et d'utiliser le principe de correspondance pour remplacer $\vec{p}$ et $\vec{r}$ par des opérateurs : $\vec{p} \to \vec{P} = -\mathrm{i}\hbar\vec{\nabla}, \vec{r} \to \vec{R}$.

Si φ est solution de l'équation de Schrödinger dans le potentiel $(\vec{A}, \overline{V})$, alors φ' en sera solution dans le potentiel transformé de jauge $(\vec{A}', \overline{V}')$ (10.58) où

$$\overline{V}' = \overline{V} + \frac{\partial\Lambda}{\partial t}$$

lorsque Λ dépend du temps. En effet, l'équation de Schrödinger pour φ s'écrit

$$\mathrm{i}\hbar\frac{\partial\varphi}{\partial t} = \frac{1}{2m}(-\mathrm{i}\hbar\vec{D})^2\varphi + q\overline{V}\varphi$$

5. Nous avons montré au § 7.4.2 que l'interaction $W = -\gamma\vec{S}\cdot\vec{B}$ entre un moment magnétique de spin et un champ magnétique découlait du couplage minimal appliqué à une équation d'onde invariante de Galilée, avec $\gamma = q_e/m_e$. Cette interaction se déduit aussi de l'équation relativiste de Dirac et de l'application de la prescription du couplage minimal à cette équation, qui donne le même le facteur gyromagnétique γ (§ 19.3.3). Les corrections de type moment magnétique anormal se déduisent du couplage minimal appliqué à l'électrodynamique quantique.

Mais, d'une part,

$$\frac{\partial\varphi}{\partial t} = \frac{\partial}{\partial t}\left(\exp\left(\frac{\mathrm{i}q\Lambda}{\hbar}\right)\varphi'\right) = \Omega^{-1}\left(\frac{\mathrm{i}q}{\hbar}\frac{\partial\Lambda}{\partial t}\varphi' + \frac{\partial\varphi'}{\partial t}\right)$$

et, d'autre part,

$$\frac{1}{2m}(-\mathrm{i}\hbar\vec{D})^2\varphi = \frac{1}{2m}(-\mathrm{i}\hbar\vec{D})^2\Omega^{-1}\varphi' = \Omega^{-1}\frac{1}{2m}(-\mathrm{i}\hbar\vec{D}')^2\varphi'$$

ce qui donne, en mettant Ω^{-1} en facteur dans les deux membres de l'équation de Schrödinger pour φ'

$$\mathrm{i}\hbar\frac{\partial\varphi'}{\partial t} = \frac{1}{2m}(-\mathrm{i}\hbar\vec{D}')^2\varphi' + q\overline{V}'\varphi'$$

On vérifie également (exercice 10.4.6) que $\vec{\jmath}$ obéit à l'équation de continuité

$$\frac{\partial|\varphi|^2}{\partial t} + \vec{\nabla}\cdot\vec{\jmath} = 0 \qquad (10.62)$$

10.3.2 Champ magnétique uniforme : niveaux de Landau

Comme application, étudions le mouvement d'une particule chargée dans un champ magnétique uniforme et constant. Nous ignorons les effets de spin : l'interaction d'un moment magnétique lié au spin a déjà été étudiée au § 3.2.5. Nous supposons $\vec{B}$ dirigé suivant Oz, et afin de simplifier la discussion, nous supposons également que le mouvement est confiné au plan xOy. Ce cas est d'ailleurs d'un grand intérêt pratique, car on sait réaliser des structures bidimensionnelles, avec des applications importantes comme l'effet Hall quantique[6]. Le mouvement d'une particule classique sous l'action de la force

$$\vec{F} = q\vec{v}\times\vec{B}$$

se fait alors sur un cercle[7] de rayon $\rho = mv/(|q|B)$ avec une fréquence $\omega = |q|B/m$, ou fréquence de Larmor (*cf.* (3.61)). Si $q < 0$, ce que nous supposerons pour fixer les idées, le cercle est parcouru dans le sens trigonométrique. Le mouvement est donc

$$\begin{aligned} x(t) &= x_0 + \rho\cos\omega t \\ y(t) &= y_0 + \rho\sin\omega t \end{aligned} \qquad (10.63)$$

où x_0 et y_0 sont les coordonnées du centre du cercle. La projection de ce mouvement circulaire uniforme sur les axes Ox et Oy donne deux oscillateurs

6. *cf.* Taylor et Heinonen [2002], chapitre 10.

7. Si le mouvement est à trois dimensions, la trajectoire est une hélice qui se projette suivant un cercle de rayon ρ parcouru à la fréquence ω dans le plan xOy.

harmoniques, que l'on retrouvera en mécanique quantique. Un choix possible pour le potentiel vecteur est

$$\vec{A} = \frac{1}{2}\,\vec{B}\times\vec{r} \tag{10.64}$$

et donc $A_x = -yB/2$, $A_y = xB/2$, $A_z = 0$. Ce choix n'est évidemment pas unique et un autre choix fréquent[8] est $A_x = A_z = 0, A_y = xB$. Calculons le commutateur des composantes de la vitesse

$$\begin{aligned}[\dot{X},\dot{Y}] &= \frac{1}{m^2}\left[P_x + \frac{q}{2}\,YB, P_y - \frac{q}{2}\,XB\right] \\ &= \frac{1}{m^2}\,\frac{qB}{2}\Big(-[P_x,X] + [Y,P_y]\Big) = -\frac{\mathrm{i}\hbar\omega}{m}I \end{aligned} \tag{10.65}$$

Comme l'expression de H peut s'écrire

$$H = \frac{1}{2}\,m\left(\dot{X}^2 + \dot{Y}^2\right) \tag{10.66}$$

on se ramène à (10.9) en définissant

$$\hat{Q} = \sqrt{\frac{m}{\hbar\omega}}\;\dot{Y} \qquad\qquad \hat{P} = \sqrt{\frac{m}{\hbar\omega}}\;\dot{X}$$

et

$$H = \frac{1}{2}\,\hbar\omega\left(\hat{P}^2 + \hat{Q}^2\right) \tag{10.67}$$

Les niveaux d'énergie sont étiquetés par un entier n

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \qquad\qquad n = 0,\ 1,\ 2,\ \ldots \tag{10.68}$$

Ces niveaux sont appelés *niveaux de Landau*. Par analogie avec le cas classique, on définit un opérateur R^2 qui est l'analogue du rayon carré ρ^2 de la trajectoire circulaire

$$R^2 = \frac{1}{\omega^2}\,\left(\dot{X}^2 + \dot{Y}^2\right) = \frac{2H}{m\omega^2} \tag{10.69}$$

La valeur moyenne de R^2 dans l'état $|n\rangle$ est

$$\langle R^2\rangle_n = \frac{2}{m\omega^2}\,\langle n|H|n\rangle = \frac{2\hbar}{m\omega}\,\left(n + \frac{1}{2}\right)$$

Si la particule est dans un état propre de H, la dispersion sur R^2 est nulle. Le flux Φ du champ magnétique à travers une orbite est quantifié en unités de $h/|q|$. En effet,

$$\Phi = \pi\langle R^2\rangle_n\,B = \frac{h}{|q|}\,\left(n + \frac{1}{2}\right)$$

8. C'est par exemple le choix de Landau et Lifschitz [1966], § 111.

La deuxième caractéristique du mouvement est la position du centre du cercle. Suivant (10.63), on définit les opérateurs X_0 et Y_0 par

$$X_0 = X - \frac{1}{\omega}\,\dot{Y} \qquad Y_0 = Y + \frac{1}{\omega}\,\dot{X} \tag{10.70}$$

Compte tenu de $[X, \dot{X}] = [Y, \dot{Y}] = \mathrm{i}\hbar I/m$ et de (10.65), le commutateur $[X_0, Y_0]$ vaut

$$[X_0, Y_0] = \frac{\mathrm{i}\hbar}{m\omega}\,I$$

On vérifie immédiatement que

$$[X_0, \dot{X}] = [X_0, \dot{Y}] = [Y_0, \dot{X}] = [Y_0, \dot{Y}] = 0$$

et donc que $[H, X_0] = [H, Y_0] = 0$. L'opérateur R_0^2

$$R_0^2 = X_0^2 + Y_0^2 \tag{10.71}$$

commute avec R^2 : R^2 et R_0^2 sont hermitiens et simultanément diagonalisables. Posant

$$\hat{Q}_0 = \sqrt{\frac{m\omega}{\hbar}}\,X_0 \qquad \hat{P}_0 = \sqrt{\frac{m\omega}{\hbar}}\,Y_0$$

conduit à

$$R_0^2 = \frac{\hbar}{m\omega}\,\left(\hat{Q}_0^2 + \hat{P}_0^2\right)$$

et les valeurs propres r_0^2 de R_0^2 sont

$$r_0^2 = \frac{2\hbar}{m\omega}\left(p + \frac{1}{2}\right) \qquad p = 0,\ 1,\ 2,\ \ldots \tag{10.72}$$

Nous avons retrouvé les deux oscillateurs harmoniques ; le premier donne la valeur n du niveau de Landau, c'est-à-dire le rayon de l'orbite, et le second la position du centre de l'orbite. Supposons que la particule se trouve dans le plan à l'intérieur d'un cercle de rayon r_0 et que $\rho^2 \ll r_0^2$; les valeurs de p sont alors limitées par

$$p \leq \frac{m\omega}{2\hbar}\,r_0^2 = \frac{m\omega}{2\pi\hbar}\,\mathcal{S}$$

où $\mathcal{S} = \pi r_0^2$ est l'aire du cercle. La dégénérescence g d'un niveau de Landau n est donnée par le nombre de valeurs possibles de p

$$g = \frac{m\omega}{2\pi\hbar}\,\mathcal{S} = \frac{|q|B}{2\pi\hbar}\,\mathcal{S} \tag{10.73}$$

Il faut multiplier ce résultat par un facteur 2 si l'on veut tenir compte du spin. En toute rigueur, il faudrait s'assurer qu'il n'y a pas de dégénérescence supplémentaire, en montrant que tout opérateur commutant avec H (ou R^2) et R_0^2 est une fonction de H et de R_0^2 : on ne peut pas trouver de propriétés

physiques supplémentaires compatibles et indépendantes. La démonstration suit le même canevas que celle donnée pour l'oscillateur harmonique simple (exercice 10.4.2).

Il n'est pas difficile de généraliser au cas d'un mouvement à trois dimensions. En effet, comme $A_z = 0$, il suffit de rajouter au hamiltonien un terme en $P_z^2/(2m)$ dont les valeurs propres sont $p_z^2/(2m)$. L'énergie totale est une fonction de n et de p_z

$$E_{n,p_z} = \hbar\omega\left(n + \frac{1}{2}\right) + \frac{p_z^2}{2m} \tag{10.74}$$

Si le mouvement vertical de la particule est limité par $0 \leq z \leq L_z$, le nombre de niveaux de Landau dans $[p_z, p_z + \Delta p_z]$ est

$$g = \frac{L_z}{2\pi\hbar}\,\frac{|q|B}{2\pi\hbar}\,\mathcal{S}\,\Delta p_z \tag{10.75}$$

10.4 Exercices

10.4.1 Éléments de matrice de Q et de P

1. Calculer les éléments de matrice $\langle n|Q|m\rangle$ et $\langle n|P|m\rangle$ des opérateurs Q et P dans la base $|n\rangle$.

2. Calculer la valeur moyenne $\langle n|Q^4|n\rangle$ de Q^4 dans l'état $|n\rangle$. Suggestion : calculer

$$|\varphi_n\rangle = \left(a + a^\dagger\right)^2 |n\rangle$$

et $||\varphi_n||^2$.

10.4.2 Propriétés mathématiques

1. Montrer les relations de commutation

$$[N, a^p] = -pa^p \quad \text{et} \quad [N, a^{\dagger p}] = pa^{\dagger p}$$

En déduire que les seules fonctions de a et de $a^\dagger$ qui commutent avec N sont des fonctions de N, et que les valeurs propres de N sont non dégénérées.

2. Soit $\mathcal{H}'$, le sous espace de $\mathcal{H}$ sous-tendu par les vecteurs $|n\rangle$ et $\mathcal{H}'_\perp$, l'espace orthogonal : $\mathcal{H} = \mathcal{H}' \oplus \mathcal{H}'_\perp$. On appelle $\mathcal{P}'$ le projecteur sur $\mathcal{H}'$. Montrer que $\mathcal{P}'$ commute avec a et $a^\dagger$ et en déduire, en utilisant le théorème de von Neumann du § 7.3.2, que soit $\mathcal{P}' = 0$, soit $\mathcal{P}' = I$. La première éventualité étant exclue, $\mathcal{P}' = I$ et les vecteurs $|n\rangle$ forment une base de $\mathcal{H}$.

10.4.3 États cohérents

1. Calculer $\langle z|P^2|z\rangle$ et $\langle z|H^2|z\rangle$ et en déduire les dispersions (10.39).

2. On se propose de rechercher les états $|\varphi(t)\rangle$ tels que les valeurs moyennes de a et de H aient des propriétés identiques aux propriétés classiques. En premier lieu, si $\langle a\rangle(t) = \langle\varphi(t)|a|\varphi(t)\rangle$, montrer que

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\langle a\rangle(t) = \omega\langle a\rangle(t)$$

et $\langle a\rangle(t)$ vérifie donc la même équation différentielle (10.29) que $z(t)$. On définit le nombre complexe z_0 par

$$z_0 = \langle a\rangle(t=0) = \langle\varphi(0)|a|\varphi(0)\rangle$$

et on aura donc comme solution de l'équation différentielle pour $\langle a\rangle(t)$

$$\langle a\rangle(t) = z_0\mathrm{e}^{-\mathrm{i}\omega t}$$

3. La deuxième condition concerne la valeur moyenne du hamiltonien. On exige, en suivant (10.30) et en ajoutant l'énergie de point zéro

$$\langle\varphi(0)|H|\varphi(0)\rangle = \hbar\omega\left(|z_0|^2 + \frac{1}{2}\right)$$

ou, de façon équivalente

$$\langle a^\dagger a\rangle = \langle\varphi(0)|a^\dagger a|\varphi(0)\rangle = |z_0|^2$$

Soit l'opérateur $b(z_0) = a - z_0$. Montrer que

$$\langle\varphi(0)|b^\dagger(z_0)b(z_0)|\varphi(0)\rangle = 0$$

et en déduire

$$a|\varphi(0)\rangle = z_0|\varphi(0)\rangle$$

L'état $|\varphi(0)\rangle$ est donc l'état cohérent $|z_0\rangle$.

4. Fonction d'onde d'un état cohérent. Exprimer l'opérateur de déplacement $D(z)$ (10.40) en fonction des opérateurs P et Q et calculer la fonction d'onde $\psi_z(q) = \langle q|z\rangle$. Suggestion : écrire $D(z)$ sous la forme

$$D(z) = f(z, z^*)\exp[c(z - z^*)Q]\,\exp[\mathrm{i}c'(z + z^*)P]$$

déterminer les constantes c et c' et utiliser le fait que P est le générateur infinitésimal des translations (*cf.* § 8.1.1)

$$\exp\left(-\mathrm{i}\frac{Pl}{\hbar}\right)|q\rangle = |q + l\rangle$$

Exprimer $\psi_z(q)$ en fonction de la fonction d'onde $\varphi_0(q)$ (10.23) de l'état fondamental.

5. Montrer qu'un opérateur A est entièrement déterminé par ses "éléments diagonaux" $\langle z|A|z\rangle$. Suggestion : utiliser

$$\langle z|A|z\rangle = \mathrm{e}^{-|z|^2}\sum_{n,m}\frac{A_{nm}z^n z^{*m}}{\sqrt{n!m!}}$$

10.4.4 Couplage à une force classique

Les états cohérents permettent de traiter simplement la version quantique de l'oscillateur harmonique forcé. En mécanique classique élémentaire, l'action d'une force externe $F(t)$ sur un oscillateur harmonique

$$m\ddot{q}(t) = -m\omega^2 q + F(t)$$

se traduit dans le hamiltonien par un couplage $-qF(t)$ entre l'élongation q et la force $F(t)$. Dans la version quantique, au hamiltonien de l'oscillateur harmonique simple (10.9) s'ajoute un terme de couplage $-QF(t)$ entre l'élongation Q et la force externe $F(t)$, que l'on appelle conventionnellement *force classique* ou aussi *source classique*

$$W(t) = -Q\sqrt{\frac{\hbar}{2m\omega}}\left(a + a^\dagger\right)F(t) = -\hbar\left(a + a^\dagger\right)f(t)$$

Dans cette expression, $f(t)$ est un nombre ayant la dimension d'une fréquence. Nous noterons H_0 le hamiltonien (10.9) de l'oscillateur harmonique simple et $H(t)$ le hamiltonien total

$$H(t) = H_0 + W(t) = H_0 - \hbar\left(a + a^\dagger\right)f(t) \tag{10.76}$$

1. Le problème est très semblable à celui rencontré au § 8.5.3 (*cf.* (8.138)), et nous pourrions tenter de le traiter par une méthode de perturbations. Cependant, il se trouve que l'on peut calculer exactement l'évolution temporelle grâce à l'opérateur de déplacement $D(z)$ (10.40). On commence par passer dans le point de vue de l'interaction (§ 4.2.5). Montrer que

$$\tilde{W}(t) = -\hbar\left(a\,\mathrm{e}^{-\mathrm{i}\omega t} + a^\dagger\,\mathrm{e}^{\mathrm{i}\omega t}\right)f(t)$$

et que l'opérateur d'évolution, dans ce point de vue, vérifie l'équation différentielle ($\tilde{U}(t) = \tilde{U}(t,0)$)

$$\mathrm{i}\,\frac{\mathrm{d}\tilde{U}(t)}{\mathrm{d}t} = -\left(a\,\mathrm{e}^{-\mathrm{i}\omega t} + a^\dagger\,\mathrm{e}^{\mathrm{i}\omega t}\right)f(t)\tilde{U}(t) \qquad \tilde{U}(t=0) = I$$

2. On divise l'intervalle $[0,t]$ en n intervalles infinitésimaux Δt et on part de

$$\tilde{U}(t) = \lim_{\Delta t\to 0}\prod_{j=0}^{n-1}\mathrm{e}^{-\mathrm{i}\Delta t\tilde{W}(t_j)/\hbar}$$

Pour calculer ce produit, on utilise la propriété (10.42) de l'opérateur de déplacement avec $z_j = \mathrm{i}f(t_j)\exp[\,\mathrm{i}\omega t_j]$. Montrer que

$$\tilde{U}(t) = \mathrm{e}^{\mathrm{i}\Phi}\exp\left(-\frac{1}{2}\int_0^t \mathrm{d}t'\mathrm{d}t''f(t')f(t'')\mathrm{e}^{-\mathrm{i}\omega(t'-t'')}\right)\exp[z(t)a^\dagger]\exp[-z^*(t)a] \tag{10.77}$$

avec

$$z(t) = \mathrm{i} \int_0^t \mathrm{d}t' \, \mathrm{e}^{\mathrm{i}\omega t'} \, f(t')$$

Le facteur de phase $\exp(\mathrm{i}\Phi)$ peut être ignoré, mais il est facile de le calculer

$$\mathrm{i}\Phi = -\frac{1}{2} \int_0^t \mathrm{d}t'\mathrm{d}t'' f(t')f(t'')\mathrm{e}^{-\mathrm{i}\omega(t'-t'')}\varepsilon(t'-t'')$$

où $\varepsilon(t)$ est la fonction signe de t : $\varepsilon(t) = 1$ si $t > 0$, $\varepsilon(t) = -1$ si $t < 0$. Vérifier que le membre de droite de cette équation est bien un nombre imaginaire pur. En déduire l'expression de $\tilde{U}(t)$ sous la forme

$$\tilde{U}(t) = \mathrm{e}^{\mathrm{i}\Phi} \, \mathrm{e}^{-|z|^2/2} \, \mathrm{e}^{a^\dagger z(t)} \, \mathrm{e}^{-az^*(t)} = \mathrm{e}^{-Y} \mathrm{e}^{a^\dagger z(t)} \, \mathrm{e}^{-az^*(t)}$$

$$Y = \int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \; \mathrm{e}^{-\mathrm{i}\omega(t'-t'')} \, f(t')f(t'')\theta(t'-t'') \qquad (10.78)$$

où $\theta(t)$ est la fonction de Heaviside : $\theta(t) = 1$ si $t > 0$, $\theta(t) = 0$ si $t < 0$. Dans la suite du problème, on se servira de la forme (10.78) de $\tilde{U}(t)$.

3. On examine le cas où l'état initial à $t = 0$ est un état propre $|n\rangle$ de H_0 en supposant que la force agit seulement pendant un intervalle de temps fini $[t_1, t_2]$, et en choisissant d'observer l'oscillateur à un temps $t > t_2$: $0 < t_1 < t_2 < t$. On définit la transformée de Fourier $\tilde{f}(\omega)$ de $f(t)$

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}t' \, \mathrm{e}^{\mathrm{i}\omega t'} f(t') = \int_{t_1}^{t_2} \mathrm{d}t' \, \mathrm{e}^{\mathrm{i}\omega t'} f(t')$$

Montrer que le résultat final pour $\tilde{U}(t)$ est indépendant de t pour $t > t_2$

$$\tilde{U}(t) = \exp(\mathrm{i}\Phi) \, \exp\left(\mathrm{i}a^\dagger \tilde{f}(\omega\right) \, \exp\left(\mathrm{i}a\tilde{f}^*(\omega\right) \, \exp\left(-\frac{1}{2}\,|\tilde{f}(\omega)|^2\right) \qquad (10.79)$$

Montrer que si l'oscillateur est dans son état fondamental au temps $t = 0$, le vecteur d'état final est un état cohérent

$$\tilde{U}(t)|0\rangle = \mathrm{e}^{\mathrm{i}\Phi} \, |\mathrm{i}\tilde{f}(\omega)\rangle$$

Montrer que la probabilité d'observer un état final $|m\rangle$ est donnée par une loi de Poisson (10.34)

$$\mathsf{p}(m) = \frac{\left(|\tilde{f}(\omega)|^2\right)^m \, \exp\left(-|\tilde{f}(\omega)|^2\right)}{m!}$$

10.4.5 Opérateur de phase

1. Afin d'essayer de définir un opérateur de phase Φ, on écrit l'opérateur d'annihilation a sous la forme d'une quasi-représentation module et phase

$$a = A\,\mathrm{e}^{\mathrm{i}\Phi} \qquad\qquad A = A^\dagger$$

En déduire $a^\dagger a + I = A^2$ et

$$A = \sqrt{N+1} = \sum_{n=0}^{\infty} |n\rangle \sqrt{n+1}\, \langle n|$$

On a donc $\exp(\mathrm{i}\Phi) = A^{-1}a$. Montrer que $\exp(\mathrm{i}\Phi)$ est identique à l'opérateur E défini en (10.47), et n'est donc pas unitaire : Φ n'existe pas comme opérateur hermitien.

2. Montrer que les opérateurs C et S définis par

$$C = \frac{1}{2}\left(E + E^\dagger\right) \qquad S = \frac{1}{2\mathrm{i}}\left(E - E^\dagger\right)$$

sont les analogues de $\cos\phi$ et $\sin\phi$. Calculer les commutateurs $[C, S]$, $[C, N]$, $[S, N]$ et la somme $C^2 + S^2$. Quelles inégalités de Heisenberg peut-on déduire de ces commutateurs ?

3. Calculer les valeurs moyennes et les dispersions de C et S dans un état cohérent $|z\rangle$, $z = |z| \exp(\mathrm{i}\theta)$.

10.4.6 Conservation du courant en présence d'un champ magnétique

En utilisant l'équation de Schrödinger dans un champ magnétique, montrer que le courant $\vec{\jmath}$ (10.59) obéit à l'équation de continuité

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{\jmath} = 0$$

10.4.7 Transformations de jauge non abéliennes

Les interactions fondamentales (à l'exception de la gravitation) sont toutes fondées sur les théories de jauge non abéliennes, que nous allons définir dans un cas élémentaire en généralisant la transformation de jauge (10.58). En omettant la dépendance par rapport au temps afin de simplifier la discussion, nous allons supposer que la fonction d'onde $\varphi(\vec{r})$ est un vecteur à deux composantes $\Phi(\vec{r}) = [\varphi_1(\vec{r}), \varphi_2(\vec{r})]$ dans un espace de Hilbert complexe à deux dimensions, et qu'il existe dans cet espace une opération de symétrie, dite de *symétrie interne*, qui laisse la physique invariante,

$$\Phi(\vec{r}) \to \Phi'(\vec{r}) = \Omega\Phi \quad \text{ou} \quad \varphi'_\alpha = \sum_{\beta=1}^{2} \Omega_{\alpha\beta}\varphi_\beta$$

généralisant (10.55). Ω est une matrice 2×2 unitaire et de déterminant unité, ou en d'autres termes une matrice du groupe $SU(2)$. La symétrie est appelée *symétrie de jauge* et le groupe $SU(2)$ est le *groupe de jauge.* En général, le groupe de jauge est un groupe de Lie compact. Le groupe de jauge de

l'électromagnétisme est le groupe des transformations de phase (10.54), noté $U(1)$, qui est abélien : l'électromagnétisme est une théorie de jauge abélienne. Lorsque le groupe de jauge est non abélien, la théorie de jauge sera dite non abélienne. Les groupes de jauge du modèle standard de la physique des particules élémentaires sont le groupe produit direct $SU(2) \times U(1)$ pour les interactions électrofaibles et $SU(3)$ pour la chromodynamique quantique, qui sont des groupes non abéliens.
Suivant les résultats de l'exercice 3.3.7, la matrice Ω peut s'écrire en fonction des matrices de Pauli

$$\Omega = \exp\left(-\mathrm{i}\, q \sum_{a=1}^{3} \Lambda_a \frac{1}{2}\, \sigma_a\right)$$

Lorsque les fonctions Λ_a sont indépendantes de $\vec{r}$, on a affaire à une symétrie de jauge *globale*, et si les Λ_a sont fonctions de $\vec{r}$, à une symétrie de jauge *locale*. Afin d'alléger les notations, on se placera dans un système d'unités où $\hbar = m = 1$.

1. L'analogue du potentiel vecteur de l'électromagnétisme est un champ vectoriel de composantes $\vec{A}_a$ dans l'espace de symétrie interne. On définit la matrice $\vec{\mathbf{A}}$ par

$$\vec{\mathbf{A}} = \sum_{a=1}^{3} \vec{A}_a \frac{1}{2}\, \sigma_a$$

$\vec{\mathbf{A}}$ possède à la fois des composantes ordinaires $i = (x, y, z)$ et des composantes a dans l'espace de symétrie interne : $\vec{\mathbf{A}} = \{A_{ia}\}$. L'expression du courant $\vec{\jmath}$ généralise (10.59)

$$\vec{\jmath} = \mathrm{Re}\left[\Phi^\dagger(-\mathrm{i}\vec{\nabla} - q\vec{\mathbf{A}})\Phi\right] = \mathrm{Re}\left[\Phi^\dagger(-\mathrm{i}\vec{\mathbf{D}}\Phi)\right]$$

où

$$\vec{\mathbf{D}} = -\mathrm{i}\vec{\nabla} - q\vec{\mathbf{A}}$$

est la dérivée covariante. Montrer que la transformation de jauge $\Phi \to \Phi'$ laisse $\vec{\jmath}$ invariant si cette transformation de jauge est globale à condition de transformer aussi $\vec{\mathbf{A}}$ en $\vec{\mathbf{A}}'$

$$\vec{\mathbf{A}}' = \Omega\vec{\mathbf{A}}\Omega^{-1}$$

Si la transformation de jauge est locale, montrer que l'invariance du courant

$$\vec{\jmath} = \vec{\jmath}\,' = \mathrm{Re}\left[\Phi'^\dagger(-\mathrm{i}\vec{\nabla} - q\vec{\mathbf{A}}')\Phi'\right]$$

implique la loi de transformation $\vec{\mathbf{A}} \to \vec{\mathbf{A}}'$

$$\vec{\mathbf{A}}' = \Omega\vec{\mathbf{A}}\Omega^{-1} - \frac{\mathrm{i}}{q}\,(\vec{\nabla}\Omega)\Omega^{-1}$$

Retrouver la loi de transformation (10.58) dans le cas abélien.

2. On choisit une transformation de jauge infinitésimale : $|q\Lambda_a(\vec{r})| \ll 1$. En déduire la loi de transformation de $\vec{A}_a$

$$\delta\vec{A}_a = \vec{A}'_a - \vec{A}_a = -\vec{\nabla}\Lambda_a + q\sum_{b,c}\varepsilon_{abc}\Lambda_b\vec{A}_c$$

La différence (cruciale !) avec le cas abélien est que le *champ de jauge* $\vec{A}_a$ dépend de façon non triviale de l'indice de symétrie interne a du groupe de jauge[9]. En électromagnétisme, les photons ne transportent pas de charge, mais les bosons de jauge d'une théorie non abélienne sont “chargés” : ils transportent les nombres quantiques de la symétrie interne.

3. Montrer que si Φ obéit à l'équation de Schrödinger indépendante du temps

$$\frac{1}{2}\left(-\mathrm{i}\nabla - q\vec{\mathbf{A}}\right)^2\Phi = \frac{1}{2}(-\mathrm{i}\hbar\vec{\mathbf{D}})^2\Phi = E\Phi$$

il en est alors de même pour Φ' à condition d'utiliser le champ $\vec{\mathbf{A}}'$.

10.5 Bibliographie

La diagonalisation du hamiltonien de l'oscillateur harmonique à une dimension par une méthode algébrique est classique et se trouve dans tous les manuels de mécanique quantique. La théorie des états cohérents est développée par Cohen-Tannoudji *et al.* [1973], complément $\mathrm{G_V}$ ou Gerry et Knight [2005], chapitre 3. Feynman *et al.* [1965], volume III, chapitre 21 donnent une discussion physique de la différence entre la vitesse et $\vec{p}/m$ en présence d'un champ électromagnétique. Les niveaux de Landau sont traités dans Cohen-Tannoudji *et al.* [1973], complément $\mathrm{E_{VI}}$ et les applications à la physique du solide par Huang [1963] chapitre 11.

9. Comme le champ $\vec{A}$ est un champ vectoriel, les particules associées sont, comme le photon, des particules de spin 1, appelées *bosons de jauge* : bosons Z^0 et $W^\pm$ pour les interactions électrofaibles, gluons pour la chromodynamique.

Chapitre 11

Intrication et non localité quantiques

Dans la section 9.6, nous avons abordé les états à deux particules et nous y revenons de façon systématique dans ce chapitre. Une fois le cas à deux particules assimilé, il n'est pas difficile de généraliser à un nombre quelconque de particules. Les états à deux particules (ou plus) conduisent à des configurations très riches, dites intriquées. Ces configurations sont telles qu'aucune distribution de probabilité classique n'est capable de reproduire leurs corrélations, *et celles-ci possèdent donc un caractère intrinsèquement quantique*. Une propriété remarquable des états intriqués est que deux systèmes quantiques peuvent rester intriqués, même s'ils sont arbitrairement éloignés, et ils continuent à former une entité indissociable. Le formalisme du produit tensoriel développé dans la section 2.4 nous permettra dans la section 11.1 de décrire les mélanges quantiques à l'aide du formalisme de l'opérateur statistique. Les sections 11.2 et 11.3 seront consacrées aux corrélations quantiques et à leurs applications aux inégalités de Bell. La section 11.4 contient une introduction à la théorie de la décohérence et à celle de la mesure. Enfin, dans la dernière section, nous passerons brièvement en revue les applications potentielles à l'information quantique. Ce dernier sujet, actuellement en pleine expansion, trouve ses applications dans le calcul, la cryptographie et la téléportation quantiques.

11.1 Opérateur statistique (ou opérateur densité)

11.1.1 Définition et propriétés

Considérons un système de deux particules[1] $\mathcal{AB}$ décrit par un vecteur d'état $|\Phi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$; comme dans la section 2.4, $\mathrm{d_A}$ et $\mathrm{d_B}$ désignent les dimensions de $\mathcal{H}_A$ et $\mathcal{H}_B$. Si $|\Phi_{AB}\rangle$ est un produit tensoriel $|\varphi_A \otimes \varphi_B\rangle$,

1. Ce qui suit est également valable si $\mathcal{A}$ et $\mathcal{B}$ sont deux sous-systèmes. Nous avons pris le cas de deux particules pour fixer les idées.

le vecteur d'état de la particule $\mathcal{A}$ est $|\varphi_A\rangle$. Mais que se passe-t-il si $|\Phi_{AB}\rangle$ n'est pas un produit tensoriel, ou, en d'autres termes, si $|\Phi_{AB}\rangle$ est un état intriqué comme l'état $|\Phi\rangle$ (2.63) ? Peut-on encore parler du vecteur d'état de la particule $\mathcal{A}$? Nous allons voir que la réponse à cette question est *négative* : en général, on ne peut pas attribuer un vecteur d'état à la particule $\mathcal{A}$. Cet exemple montre qu'il faut généraliser notre description des systèmes quantiques. Si l'état physique de la particule $\mathcal{A}$ peut être décrit par un vecteur de l'espace de Hilbert, on dit que l'on a affaire à un *état pur* ou à un *cas pur* ; ceci sera le cas si l'on dispose d'une information maximale sur la particule $\mathcal{A}$. Lorsque l'information sur $\mathcal{A}$ est incomplète[2], on a affaire à un *mélange*, et l'état quantique de la particule $\mathcal{A}$ est alors décrit mathématiquement par un opérateur statistique[3] : le système $\mathcal{AB}$ est dans un état pur $|\Phi_{AB}\rangle$, mais la particule $\mathcal{A}$ ne l'est pas. L'introduction de l'opérateur statistique va nous permettre de reformuler le postulat **I** du chapitre 4 afin de décrire des situations physiques plus générales que celles envisagées jusqu'ici, celles où l'on ne dispose que d'une information partielle sur le système considéré.

Lorsque l'on a affaire à un cas pur, la donnée du vecteur d'état $|\varphi\rangle \in \mathcal{H}$ pour décrire un système quantique est équivalente à celle du projecteur $\mathcal{P}_\varphi = |\varphi\rangle\langle\varphi|$ sur l'état $|\varphi\rangle$. En un certain sens, se donner $\mathcal{P}_\varphi$ est même préférable, car la phase arbitraire de $|\varphi\rangle$ disparaît : $\mathcal{P}_\varphi$ est invariant lorsque l'on multiplie $|\varphi\rangle$ par un facteur de phase

$$|\varphi\rangle \to \mathrm{e}^{\,\mathrm{i}\alpha}\,|\varphi\rangle$$

et il y a donc correspondance biunivoque entre l'état physique et $\mathcal{P}_\varphi$, et non pas correspondance à un facteur de phase près. La valeur moyenne d'une propriété physique A s'exprime simplement en fonction de $\mathcal{P}_\varphi$. Introduisons en effet une base orthonormée $|n\rangle$ de $\mathcal{H}$

$$\begin{aligned} \langle A\rangle &= \langle\varphi|A|\varphi\rangle = \sum_{n,m}\langle\varphi|n\rangle\langle n|A|m\rangle\langle m|\varphi\rangle \\ &= \sum_{n,m}\langle m|\varphi\rangle\langle\varphi|n\rangle\langle n|A|m\rangle \\ &= \sum_{m}\langle m|\mathcal{P}_\varphi\,A|m\rangle = \mathrm{Tr}(\mathcal{P}_\varphi A) \end{aligned} \tag{11.1}$$

$\mathcal{P}_\varphi$ est le cas le plus simple d'opérateur statistique. Nous pouvons maintenant généraliser à un *mélange* : dans un tel cas, on sait seulement que le système

2. Cette formulation n'est pas satisfaisante car elle n'est pas intrinsèque : elle fait intervenir l'information que *nous* avons sur le système. Une formulation plus satisfaisante est fondée sur les corrélations internes et externes du système, voir Mermin [1998].

3. Une fois de plus, la terminologie usuelle "opérateur densité" est particulièrement malvenue. De quelle densité s'agit-il ? En fait, la terminologie a été introduite dans le cas de la mécanique ondulatoire (chapitre 8), où les éléments de matrice diagonaux de ρ dans la base des états propres de l'opérateur position, $\langle\vec{r}\,|\rho|\vec{r}\,\rangle$, ou de l'opérateur impulsion, $\langle\vec{p}\,|\rho|\vec{p}\,\rangle$, sont effectivement des densités. Cependant, "opérateur densité" masque le fait que cet opérateur contient des informations essentielles sur les phases. Nous avons préféré "opérateur statistique" ; par analogie avec vecteur d'état, "opérateur d'état" serait aussi possible.

quantique a la probabilité p_α ($0 \leq \mathsf{p}_\alpha \leq 1, \sum_\alpha \mathsf{p}_\alpha = 1$) de se trouver dans l'état $|\varphi_\alpha\rangle$. Les états $|\varphi_\alpha\rangle$ sont supposés normalisés ($\langle\varphi_\alpha|\varphi_\alpha\rangle = 1$), mais pas nécessairement orthogonaux, $\langle\varphi_\alpha|\varphi_\beta\rangle \neq \delta_{\alpha\beta}$. Par définition, *l'opérateur statistique* ρ décrivant ce système quantique est la combinaison linéaire de projecteurs $\mathcal{P}_{\varphi_\alpha}$

$$\boxed{\rho = \sum_\alpha \mathsf{p}_\alpha |\varphi_\alpha\rangle\langle\varphi_\alpha| = \sum_\alpha \mathsf{p}_\alpha \mathcal{P}_{\varphi_\alpha}} \tag{11.2}$$

Cette définition assure que la valeur moyenne d'une propriété physique A s'obtient par une généralisation immédiate de (11.2). En effet, la valeur moyenne de A dans l'état $|\varphi_\alpha\rangle$, $\langle A\rangle_\alpha$, est

$$\langle A\rangle_\alpha = \langle\varphi_\alpha|A|\varphi_\alpha\rangle$$

et cette valeur moyenne est affectée du poids p_α dans le calcul de la valeur moyenne globale $\langle A\rangle$. La valeur moyenne dans le mélange est donc

$$\boxed{\langle A\rangle = \sum_\alpha \mathsf{p}_\alpha \langle A\rangle_\alpha = \sum_\alpha \mathsf{p}_\alpha \mathrm{Tr}(\mathcal{P}_{\varphi_\alpha} A) = \mathrm{Tr}(\rho A)} \tag{11.3}$$

Les poids p_α sont fixés par le problème physique considéré, et (11.2) définit la *préparation* du système : chaque état $|\varphi_\alpha\rangle$ apparaît avec la probabilité p_α. Un cas classique est celui de la mécanique statistique d'équilibre où un système $\mathcal{A}$ est en équilibre avec un thermostat à la température absolue T (ensemble canonique) : si les niveaux d'énergie du système sont E_r, l'opérateur statistique est alors[4]

$$\rho_{\mathrm{can}} = \frac{1}{Z}\sum_r \mathrm{e}^{-E_r/k_{\mathrm{B}}T} \qquad\qquad Z = \sum_r \mathrm{e}^{-E_r/k_{\mathrm{B}}T} \tag{11.4}$$

Les propriétés fondamentales de ρ, qui suivent immédiatement de sa définition (11.2) sont :

- ρ est hermitien : $\rho = \rho^\dagger$.
- ρ est de trace unité : $\mathrm{Tr}\,\rho = 1$.
- ρ est un opérateur positif[5] : $\langle\psi|\rho|\psi\rangle \geq 0$ quel que soit $|\psi\rangle$, car $\langle\psi|\mathcal{P}_\varphi|\psi\rangle = |\langle\psi|\varphi\rangle|^2$.
- Une condition nécessaire et suffisante pour que ρ décrive un état pur est $\rho^2 = \rho$: en effet, comme $\rho = \rho^\dagger$, la condition $\rho^2 = \rho$ implique que ρ est un projecteur. Comme $\mathrm{Tr}\,\rho = 1$, la dimension de l'espace vectoriel sur lequel projette ρ est 1[6], et ρ est de la forme $|\varphi\rangle\langle\varphi|$. Une autre condition nécessaire et suffisante pour que ρ représente un cas pur est $\mathrm{Tr}\,\rho^2 = 1$ (exercice 11.6.1).

---

4. La somme porte sur les niveaux, pas sur les valeurs de l'énergie : si un niveau d'énergie est doublement dégénéré, il compte deux fois.

5. Un opérateur (strictement) positif est hermitien et a des valeurs propres (strictement) positives et réciproquement : exercice 2.4.10.

6. En général, si $\mathcal{P}$ est un projecteur, $\mathrm{Tr}\,\mathcal{P}$ est égal à la dimension de l'espace vectoriel de projection : il suffit pour le voir de se placer dans une base où $\mathcal{P}$ est diagonal.

Inversement, un opérateur hermitien, positif et de trace unité peut être interprété comme un opérateur statistique. En effet, comme ρ est hermitien, on peut l'écrire sous forme diagonale (mais la préparation n'est pas unique si certaines valeurs propres sont dégénérées)

$$\rho = \sum_i \mathsf{p}_i |i\rangle\langle i| \tag{11.5}$$

et une préparation possible du système quantique consiste à fabriquer un mélange d'états $|i\rangle$ avec une probabilité p_i. Cependant, si la donnée des p_α et des $|\varphi_\alpha\rangle$ dans (11.2) détermine ρ de façon unique, l'inverse n'est pas vrai : *à un même opérateur statistique peuvent correspondre plusieurs préparations différentes*. Par exemple, si une valeur propre p_i dans (11.5) est dégénérée, nous pouvons choisir pour la préparation des combinaisons linéaires arbitraires dans le sous-espace de cette valeur propre. En d'autres termes, un opérateur statistique ne spécifie pas une configuration microscopique unique, mais il est suffisant pour calculer les valeurs moyennes de propriétés physiques à l'aide de (11.3). Nous reviendrons au § 11.1.4 sur ce point fondamental. Notons enfin deux définitions issues de la physique atomique : dans une base où ρ n'est pas diagonal, les éléments diagonaux ρ_{ii} sont appelés *populations* et les éléments non diagonaux $\rho_{ij}, i \neq j$ sont appelés *cohérences*. Cette définition dépend bien sûr du choix de la base, mais il existe souvent un choix naturel pour une base privilégiée, par exemple la base $\{|+\rangle, |-\rangle\}$ dans le cas de la RMN pour un champ $\vec{B}_0$ parallèle à Oz, ou la base des états fondamentaux et excités dans le cas de l'atome à deux niveaux.

11.1.2 Opérateur statistique réduit

Nous avons obtenu l'opérateur statistique ρ (11.2) en construisant un mélange d'états $|\varphi_\alpha\rangle$ avec des poids p_α. Nous allons maintenant donner une seconde interprétation de l'opérateur statistique en essayant de répondre à la question suivante : supposons que le vecteur d'état d'un système $\mathcal{AB}$ soit un état intriqué $|\Phi_{AB}\rangle$, mais que nous ayons seulement la possibilité de mesurer les propriétés physiques de $\mathcal{A}$. Comment décrire l'état quantique du seul système $\mathcal{A}$? Soit donc $|\Phi_{AB}\rangle$ l'état intriqué le plus général, écrit dans la base[7] $|i \otimes m\rangle$ de $\mathcal{H}_A \otimes \mathcal{H}_B$

$$|\Phi_{AB}\rangle = \sum_{i,m} c_{im} |i \otimes m\rangle \qquad \sum_{i,m} |c_{im}|^2 = 1 \tag{11.6}$$

Soit A, une propriété physique du système $\mathcal{A}$, qui est représentée par $A \otimes I_B$ dans $\mathcal{H}_A \otimes \mathcal{H}_B$; écrivons son action sur $|\Phi_{AB}\rangle$

$$(A \otimes I_B)|\Phi_{AB}\rangle = \sum_{i,m} c_{im} |Ai \otimes m\rangle$$

7. Afin d'alléger les notations, nous écrivons $|i \otimes m\rangle$ au lieu de $|i_A \otimes m_B\rangle$.

et calculons la valeur moyenne de A

$$\begin{aligned}\langle\Phi_{AB}|A\otimes I_B|\Phi_{AB}\rangle &= \sum_{j,n}\sum_{i,m} c^*_{jn}c_{im}\langle j\otimes n|Ai\otimes m\rangle \\ &= \sum_{j,n}\sum_{i,m} c^*_{jn}c_{im}\langle j|A|i\rangle\delta_{nm} = \sum_{i,j}\sum_{m} c^*_{jm}c_{im}\,\langle j|A|i\rangle \\ &= \sum_{i,j}\rho_{ij}\,\langle j|A|i\rangle = \sum_{i,j}\rho_{ij}\,A_{ji} = \mathrm{Tr}\,(\rho A)\ , \qquad (11.7)\end{aligned}$$

où nous avons introduit l'opérateur ρ agissant dans $\mathcal{H}_A$ dont les éléments de matrice ρ_{ij} sont

$$\boxed{\rho_{ij} = \sum_m c_{im}c^*_{jm}} \qquad (11.8)$$

La notation ρ n'a pas été choisie par hasard : en effet, ρ possède toutes les propriétés d'un opérateur statistique ce que l'on peut aisément vérifier à partir de (11.8) (exercice 11.6.3) :

- ρ est hermitien : $\rho = \rho^\dagger$;
- il est de trace unité : $\mathrm{Tr}\,\rho = 1$;
- il est positif : $\forall|\psi\rangle\ \langle\psi|\rho|\psi\rangle \geq 0$.

Par construction, la valeur moyenne d'une propriété physique de $\mathcal{A}$ est donnée par $\langle A\rangle = \mathrm{Tr}\,(\rho A)$. Sauf dans le cas où $|\Phi_{AB}\rangle$ est un produit tensoriel, $|\Phi_{AB}\rangle = |\varphi_A\otimes\varphi_B\rangle$, on ne peut pas attribuer un vecteur d'état au système $\mathcal{A}$, mais seulement un opérateur statistique. Une question importante se pose alors : un opérateur statistique construit à partir de (11.3) et un opérateur statistique construit à partir de (11.8) ont-ils une interprétation physique identique, ou bien doit-on faire une distinction ? Cette question sera discutée dans la section 11.1.4.

Il est facile de généraliser (11.8) au cas où le sytème $\mathcal{AB}$ lui-même est décrit par un opérateur statistique ρ_{AB}, et non un vecteur d'état $|\Phi_{AB}\rangle$. Il suffit comme ci-dessus de calculer la valeur moyenne d'une propriété physique A, représentée par $A\otimes I_B$ dans[8] $\mathcal{H}_A\otimes\mathcal{H}_B$

$$\begin{aligned}\langle A\otimes I_B\rangle &= \mathrm{Tr}_{\mathcal{AB}}\big(\rho_{AB}[A\otimes I_B]\big) \\ &= \sum_{ijmn}\rho^{AB}_{im;jn}A_{ji}\delta_{nm} = \sum_{i,j}A_{ji}\sum_m \rho^{AB}_{im;jm} = \mathrm{Tr}(A\rho_A) \qquad (11.9)\end{aligned}$$

L'expression qui généralise (11.8) et qui donne l'opérateur statistique de $\mathcal{A}$ est donc

$$\boxed{\rho^A_{ij} = \sum_m \rho^{AB}_{im;jm} \qquad \rho_A = \mathrm{Tr}_{\mathcal{B}}\rho_{AB}} \qquad (11.10)$$

8. Afin de rendre les notations plus claires, nous avons écrit les indices A et B en exposant dans la forme matricielle de (11.9).

Il est important de noter que *la seconde expression est indépendante de la base* : $\mathrm{Tr}_{\mathcal{B}}$ représente la trace sur l'espace $\mathcal{H}_B$, appelée *trace partielle* de l'opérateur statistique global, tandis que ρ_A est *l'opérateur statistique réduit*, et on peut montrer que cet opérateur donne la solution unique de (11.9) quel que soit A (Breuer et Petruccione [2002], chapitre 2). Une application intéressante de (11.10) est le calcul de la probabilité de trouver la valeur propre a_n de A, qui est donnée en fonction du projecteur $\mathcal{P}_n$ sur le sous-espace de la valeur propre a_n d'une propriété physique A par une expression généralisant (4.4)

$$\boxed{\mathsf{p}(a_n) = \mathrm{Tr}_{\mathcal{A}}\left(\mathcal{P}_n \rho_A\right) = \mathrm{Tr}_{\mathcal{A}}\mathrm{Tr}_{\mathcal{B}}\left[(\mathcal{P}_n \otimes I_B)\rho_{AB}\right]} \tag{11.11}$$

Il est important de se convaincre que *la prescription qui consiste à prendre la trace partielle est une conséquence du postulat* **II** (règle de Born), car l'expression donnant les valeurs moyennes découle de ce postulat.

Donnons maintenant quelques exemples de calcul explicite de l'opérateur statistique réduit lorsque l'on part d'un cas pur $|\Phi_{AB}\rangle$ dans l'espace produit tensoriel $\mathcal{H}_A \otimes \mathcal{H}_B$, construit avec des vecteurs quelconques $|\varphi_i\rangle$ et $|\chi_m\rangle$

$$|\Phi_{AB}\rangle = \sum_i \sum_m c_{im} |\varphi_i \otimes \chi_m\rangle \tag{11.12}$$

Le calcul de l'opérateur statistique réduit est immédiat si l'on observe que

$$\mathrm{Tr}\, |a\rangle\langle b| = \sum_n \langle n|a\rangle\langle b|n\rangle = \sum_n \langle b|n\rangle\langle n|a\rangle = \langle b|a\rangle \tag{11.13}$$

En écrivant l'expression explicite de $|\Phi_{AB}\rangle\langle\Phi_{AB}|$, on déduit que l'opérateur d'état réduit ρ_A dans $\mathcal{H}_A$ vaut

$$\rho_A = \mathrm{Tr}_{\mathcal{B}} |\Phi_{AB}\rangle\langle\Phi_{AB}| = \sum_{i,m,k,p} c_{im} c^*_{kp} |\varphi_i\rangle\langle\varphi_k| \langle\chi_p|\chi_m\rangle \tag{11.14}$$

Lorsque les indices i et m sont identiques, ce qui implique que $|\varphi_i\rangle$ et $\chi_m\rangle$ sont corrélés, (11.12) se simplifie[9]

$$|\Phi_{AB}\rangle = \sum_{i=1}^{d} c_i |\varphi_i \otimes \chi_i\rangle$$

et ρ_A est donné par

$$\boxed{\rho_A = \sum_i c_i c^*_j |\varphi_i\rangle\langle\varphi_j| \langle\chi_j|\chi_i\rangle} \tag{11.15}$$

9. En explicitant les indices : $i \to i_A, m \to i_B$ et

$$|\Phi_{AB}\rangle = \sum_i c_i |\varphi_{i_A} \otimes \chi_{i_B}\rangle$$

.

Un cas particulier important est celui où les états $|\chi_i\rangle$ sont orthogonaux, $\langle\chi_i|\chi_j\rangle = \delta_{ij}$. Alors, les cohérences de ρ_A s'annulent et on obtient un mélange du type (11.2)

$$\boxed{\rho_A = \sum_i |c_i|^2 \, |\varphi_i\rangle\langle\varphi_i| \quad \text{si} \quad \langle\chi_i|\chi_j\rangle = \delta_{ij}} \tag{11.16}$$

Un autre résultat utile est donné par le *théorème de purification de Schmidt*. D'après ce théorème, tout état $|\Phi_{AB}\rangle$ de $\mathcal{H}_A \otimes \mathcal{H}_B$ peut s'écrire sous la forme

$$\boxed{|\Phi_{AB}\rangle = \sum_{i=1}^{d_S} \sqrt{\mathsf{p}_i} \; |i_A \otimes i_B\rangle \qquad \mathsf{p}_i > 0 \qquad \sum_i \mathsf{p}_i = 1} \tag{11.17}$$

où les bases $|i_A\rangle$ et $|i_B\rangle$ dépendent de $|\Phi_{AB}\rangle$. Le nombre d_S ($d_S \le \text{Min}(d_A, d_B)$), est appelé *nombre de Schmidt*. Une démonstration possible de ce théorème fait appel à l'opérateur statistique réduit. Soit $|\Phi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ un état pur du système couplé $\mathcal{AB}$, et soit $\{|i_A\rangle\}$ et $\{|i_B\rangle\}$ deux bases othonormales de $\mathcal{H}_A$ et $\mathcal{H}_B$. La décomposition la plus générale de $|\Phi_{AB}\rangle$ sur la base $\{|i_A \otimes i_B\rangle\}$ de $\mathcal{H}_A \otimes \mathcal{H}_B$ s'écrit

$$|\Phi_{AB}\rangle = \sum_{i_A, i_B} c_{i_A i_B} |i_A \otimes i_B\rangle$$

Définissant les vecteurs $|\tilde{i}_B\rangle \in \mathcal{H}_B$ par

$$|\tilde{i}_B\rangle = \sum_{i_B} c_{i_A i_B} |i_B\rangle,$$

nous pouvons récrire

$$|\Phi_{AB}\rangle = \sum_i |i_A \otimes \tilde{i}_B\rangle$$

Il n'y a évidemment aucune raison pour que l'ensemble $\{|\tilde{i}_B\rangle\}$ forme une base orthonormée de $\mathcal{H}_B$. Choisissons, cependant, une base *particulière* $\{|i_A\rangle\}$ de $\mathcal{H}_A$ telle que l'opérateur statistique réduit ρ_A soit diagonal

$$\rho_A = \text{Tr}_{\mathcal{B}} |\Phi_{AB}\rangle\langle\Phi_{AB}| = \sum_{i_A=1}^{d_S} \mathsf{p}_i |i_A\rangle\langle i_A|$$

Si le nombre d_S de coefficients non nuls p_i est inférieur à la dimension d_A de $\mathcal{H}_A$, on complète $\{|i_A\rangle\}$ par un ensemble de $(d_A - d_S)$ vecteurs orthonormés orthogonaux au sous-espace sous-tendu par $\{|i_A\rangle\}$. On utilise (11.16) pour calculer ρ_A

$$\rho_A = \sum_{i,j} \langle\tilde{j}_B|\tilde{i}_B\rangle |i_A\rangle\langle j_A|$$

En comparant les deux expressions de ρ_A, on constate que

$$\langle \tilde{j}_B | \tilde{i}_B \rangle = \mathsf{p}_i \delta_{ij}$$

et avec notre choix de base, il s'avère que les vecteurs $\{|\tilde{i}_B\rangle\}$ sont, tout compte fait, orthogonaux. Pour obtenir une base orthonormée, il suffit de redéfinir

$$|i_B\rangle = \mathsf{p}_i^{-1/2} |\tilde{i}_B\rangle$$

où nous pouvons supposer $\mathsf{p}_i > 0$, car, comme expliqué ci-dessus, il est toujours possible de compléter $\mathcal{H}_B$ par un ensemble de $(d_B - d_S)$ vecteurs orthonormaux. On en déduit la représentation de Schmidt (11.17) de $|\Phi_{AB}\rangle$. Le calcul des opérateurs statistiques réduits est immédiat à partir de (11.16) et (11.17)

$$\rho_A = \sum_i \mathsf{p}_i |i_A\rangle\langle i_A| \qquad \rho_B = \sum_i \mathsf{p}_i |i_B\rangle\langle i_B| \tag{11.18}$$

Les probabilités p_i sont identiques pour ρ_A et ρ_B.

Illustrons ces résultats pour un système de deux spins 1/2 dans l'état pur (2.63) que nous réécrivons ci-dessous

$$|\Phi\rangle = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle)$$

Si le spin $\mathcal{A}$ était dans un état pur, $\mathrm{Tr}_B \rho$ devrait être un projecteur. Or ceci n'est pas le cas comme on le vérifie explicitement en calculant ρ_A à l'aide de (11.16)

$$\rho_A = \mathrm{Tr}_{\mathcal{B}} \rho_{AB} = \frac{1}{2}|+\rangle\langle+| + \frac{1}{2}|-\rangle\langle-| = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \tag{11.19}$$

Même si le système des deux spins est un état pur, l'état d'un spin individuel est en général un mélange. En fait, l'opérateur statistique (11.19) constitue un cas extrême de mélange, qui correspond à un désordre maximal et à une information minimale sur le spin. On montre qu'une mesure quantitative de l'information contenue dans l'opérateur statistique est donnée par l'*entropie de von Neumann* ou *entropie statistique*[10] $S_{\mathrm{vN}} = -\mathrm{Tr}\,\rho \ln \rho$, qui est d'autant plus grande que l'information est réduite. Dans le cas d'un spin 1/2, elle est comprise entre 0 et $\ln 2$ et vaut 0 pour un cas pur, $\ln 2$ pour le mélange (11.19) : $\ln 2$ est la valeur maximale de l'entropie de von Neumann pour un spin 1/2, et le mélange (11.19) est bien celui qui contient l'information minimale. Si l'espace de Hilbert des états d'un système quantique est de dimension d, l'opérateur statistique correspondant au désordre maximal est $\rho = I/d$, soit une entropie de von Neumann $S_{\mathrm{vN}} = \ln d$.

10. Il faut prendre garde au fait que $\mathrm{Tr}\,\rho \ln \rho \neq \sum_\alpha \mathsf{p}_\alpha \ln \mathsf{p}_\alpha$, sauf si les vecteurs $|\varphi_\alpha\rangle$ dans (11.2) sont orthogonaux entre eux.

Il est intéressant de remarquer que l'opérateur statistique donne la description la plus générale possible d'un système quantique si l'on pose *a priori* un petit nombre de principes généraux. En effet, Gleason a démontré le théorème suivant :
Théorème de Gleason. Soit un ensemble de projecteurs $\mathcal{P}_i$ agissant dans un espace de Hilbert $\mathcal{H}$ et soit un test associé à chacun des $\mathcal{P}_i$. La probabilité $\mathsf{p}(\mathcal{P}_i)$ de réussir le test doit obéir aux conditions

$$0 \leq \mathsf{p}(\mathcal{P}_i) \leq 1, \qquad \mathsf{p}(I) = 1,$$

ainsi qu'à

$$\mathsf{p}(\mathcal{P}_i + \mathcal{P}_j) = \mathsf{p}(\mathcal{P}_i) + \mathsf{p}(\mathcal{P}_j) \ \text{ si } \ \mathcal{P}_i\mathcal{P}_j = \delta_{ij}\mathcal{P}_i$$

pour tout ensemble de $\mathcal{P}_i$ vérifiant $\sum_i \mathcal{P}_i = I$ et $\mathcal{P}_i\mathcal{P}_j = \delta_{ij}\mathcal{P}_i$, c'est-à-dire pour toute décomposition de $\mathcal{H}$ en sous-espaces orthogonaux. Alors, *si la dimension de* $\mathcal{H} \geq 3$, il existe un opérateur hermitien et positif ρ de trace unité tel que

$$\mathsf{p}(\mathcal{P}_i) = \mathrm{Tr}\,(\rho\mathcal{P}_i)$$

En d'autres termes, si on associe à des projecteurs des tests avec des propriétés "raisonnables", ces probabilités doivent être données par un opérateur statistique. L'hypothèse forte est en fait la suivante : considérons, pour simplifier les notations, un sous-espace à deux dimensions de $\mathcal{H}$ et deux bases orthonormées différentes $\{|i\rangle, |j\rangle\}$ et $\{|\alpha\rangle, |\beta\rangle\}$ qui définissent les projecteurs

$$\mathcal{P}_i = |i\rangle\langle i| \qquad \mathcal{P}_j = |j\rangle\langle j| \qquad \mathcal{P}_\alpha = |\alpha\rangle\langle\alpha| \qquad \mathcal{P}_\beta = |\beta\rangle\langle\beta|$$

On a, bien sûr

$$\mathcal{P}_i + \mathcal{P}_j = \mathcal{P}_\alpha + \mathcal{P}_\beta$$

En mécanique quantique, ceci entraîne immédiatement l'égalité

$$\mathsf{p}(\mathcal{P}_i) + \mathsf{p}(\mathcal{P}_j) = \mathsf{p}(\mathcal{P}_\alpha) + \mathsf{p}(\mathcal{P}_\beta)$$

mais il n'y a aucune raison *a priori* pour qu'une telle égalité soit valable dans toute théorie imaginable.

11.1.3 Opérateur statistique pour un système à deux niveaux

Comme exemple, nous allons donner la forme la plus générale de l'opérateur statistique d'un système quantique à deux niveaux : l'espace de Hilbert des états est de dimension deux. Les applications sont multiples : description de la polarisation d'une particule massive de spin 1/2, de la polarisation d'un photon, d'un atome à deux niveaux, etc. Comme un cas très courant de système à deux niveaux est celui du spin 1/2, nous utiliserons ce cas particulier pour fixer les notations et la terminologie. Choisissons deux vecteurs de base

de l'espace des états notés $|+\rangle$ et $|-\rangle$, qui sont, par exemple, les vecteurs propres de la composante z du spin. Dans cette base, l'opérateur statistique est représenté par une matrice 2×2, la *matrice statistique* ρ. Écrivons d'abord que cette matrice est hermitienne et de trace unité : la matrice la plus générale de ce type est

$$\rho = \begin{pmatrix} a & c \\ c^* & 1-a \end{pmatrix} \tag{11.20}$$

où a est un nombre réel et c un nombre complexe. L'équation (11.20) ne définit pas encore une matrice statistique, car en plus ρ doit être positive. Les valeurs propres λ_+ et λ_- de ρ vérifient

$$\lambda_+ + \lambda_- = 1 \qquad \lambda_+\lambda_- = \det\rho = a(1-a) - |c|^2$$

et on doit avoir $\lambda_+ \geq 0$ et $\lambda_- \geq 0$. La condition $\det\rho \geq 0$ implique que λ_+ et λ_- sont de même signe, et la condition $\lambda_+ + \lambda_- = 1$ que la valeur maximale du produit $\lambda_+\lambda_-$ est $1/4$, soit finalement

$$0 \leq a(1-a) - |c|^2 \leq \frac{1}{4} \tag{11.21}$$

Une condition nécessaire et suffisante pour que ρ décrive un cas pur est

$$\det\rho = a(1-a) - |c|^2 = 0$$

À titre d'exercice, on peut calculer a et c pour la matrice statistique décrivant le vecteur d'état normalisé $|\psi\rangle = \lambda|+\rangle + \mu|-\rangle$ avec $|\lambda|^2 + |\mu|^2 = 1$, et vérifier que le déterminant de la matrice s'annule.

Il est souvent commode de décomposer la matrice statistique (11.20) sur la base des matrices de Pauli σ_i ; en effet, toute matrice 2×2 hermitienne peut s'écrire comme une combinaison linéaire avec des coefficients réels de la matrice unité I et des σ_i

$$\rho = \frac{1}{2}\begin{pmatrix} 1+b_z & b_x - \mathrm{i}b_y \\ b_x + \mathrm{i}b_y & 1-b_z \end{pmatrix} = \frac{1}{2}\left(I + \vec{b}\cdot\vec{\sigma}\right) \tag{11.22}$$

Le vecteur $\vec{b}$, appelé *vecteur de Bloch*, doit vérifier $|\vec{b}|^2 \leq 1$ en raison de (11.21). Le cas pur, qui correspond à $|\vec{b}|^2 = 1$, est aussi appelé *complètement polarisé*, le cas $\vec{b} = 0$ *non polarisé*, ou de polarisation nulle, et le cas $0 < |\vec{b}| < 1$, partiellement polarisé. Pour interpréter physiquement le vecteur $\vec{b}$, on calcule la valeur moyenne du spin $\vec{S} = \frac{1}{2}\hbar\vec{\sigma}$ en utilisant $\mathrm{Tr}\,\sigma_i\sigma_j = 2\delta_{ij}$ et on en déduit

$$\langle S_i\rangle = \mathrm{Tr}\,(\rho\, S_i) = \frac{1}{2}\hbar b_i \tag{11.23}$$

$\hbar\vec{b}/2$ est donc la valeur moyenne $\langle\vec{S}\rangle$ du spin. Il est utile de donner une représentation géométrique du vecteur de Bloch (figure 11.1) : $\vec{b}$ est un vecteur

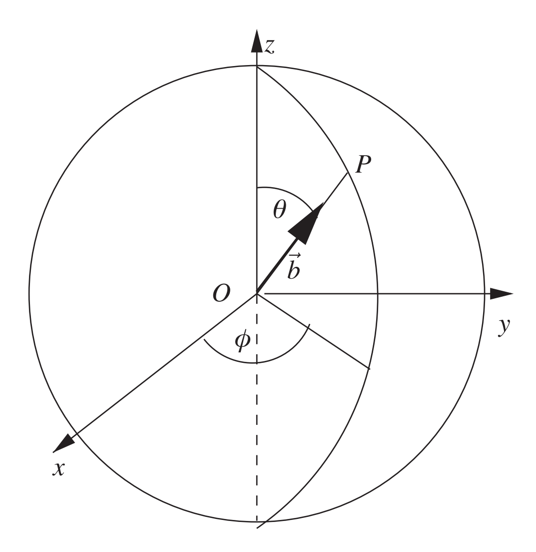


FIG. 11.1 – La sphère de Poincaré-Bloch et le vecteur de Bloch $\vec{b}$ ($OP = 1$).

d'origine O défini par ses angles polaires (θ, ϕ) et dont l'extrémité se trouve à l'intérieur d'une sphère de centre O et de rayon unité, la *sphère de Poincaré-Bloch*[11]. Un cas pur correspond à un vecteur dont l'extrémité se trouve sur la surface de la sphère.

Il est essentiel de bien faire la différence entre état pur et mélange : supposons, par exemple, qu'un spin 1/2 soit dans l'état pur

$$|\chi\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle + |-\rangle\right) \tag{11.24}$$

L'analyse par un appareil de Stern-Gerlach dont le champ magnétique $\vec{B}$ est parallèle à Oz va donner une probabilité de 50 % de déviation vers le haut et 50 % de déviation vers le bas. Cependant, l'état (11.24) est un état propre de S_x, $|\chi\rangle = |+, \hat{x}\rangle$: si $\vec{B}$ est parallèle à Ox, 100 % des spins seront déviés dans la direction des x positifs : le vecteur de Bloch est $\vec{b} = (1, 0, 0)$. Lorsque $\vec{b} = 0$, cas non polarisé où la matrice statistique est donnée par (11.19), les probabilités de déviation vers les x positifs et négatifs seront toujours de 50 % : en fait, *quelle que soit l'orientation de l'appareil de Stern-Gerlach*, il y aura toujours 50 % des spins déviés dans la direction de $\vec{B}$ et 50 % dans celle de $-\vec{B}$. La différence entre les deux cas est que pour le cas pur (11.24), état complètement polarisé, il existe une *relation de phase* bien définie entre les amplitudes pour trouver $|\chi\rangle$ dans les états $|+\rangle$ et $|-\rangle$. L'état pur $|\chi\rangle$ est une superposition *cohérente* des états $|+\rangle$ et $|-\rangle$, alors que le mélange (11.19)

11. Cette sphère a été introduite la première fois par Poincaré pour décrire la polarisation de la lumière, et redécouverte par Bloch pour la description du spin 1/2.

est une superposition *incohérente* de ces mêmes états. Dans un mélange, l'information sur les phases est perdue, au moins partiellement (car il peut bien sûr exister des états partiellement polarisés $0 < |\vec{b}| < 1$) ; elle est totalement perdue pour un état non polarisé. L'information sur les phases est contenue dans les éléments de matrice non diagonaux, les cohérences de la matrice statistique ρ.

Les mêmes remarques valent pour la polarisation de la lumière, ou la polarisation d'un photon : une lumière non polarisée est une superposition incohérente de lumière polarisée linéairement à 50 % suivant Ox et 50 % suivant Oy, sans relation de phase entre les deux. Une lumière polarisée circulairement à droite $|D\rangle$ ou à gauche $|G\rangle$ est décrite par les vecteurs (3.24)

$$|D\rangle = -\frac{1}{\sqrt{2}}\left(|x\rangle + \mathrm{i}|y\rangle\right) \qquad |G\rangle = \frac{1}{\sqrt{2}}\left(|x\rangle - \mathrm{i}|y\rangle\right) \tag{11.25}$$

Une telle lumière sera arrêtée à 50 % par un polariseur linéaire dirigé suivant Ox, et plus généralement suivant un axe quelconque, tout comme une lumière non polarisée, mais elle sera transmise à 100 % ou complètement arrêtée par un analyseur circulaire. Au contraire, si les photons sont non polarisés, tout polariseur (λ, μ) (*cf.* § 3.1.1) laissera passer les photons avec une probabilité de 50 %.

De façon générale, *la caractéristique d'un état pur est qu'il existe un test maximal dont un des résultats a une probabilité de 100 %*, tandis que pour un mélange, il n'existe pas de test maximal possédant cette propriété (exercice 11.6.1). Dans le cas du spin 1/2, cela veut dire qu'il n'existe pas d'orientation de $\vec{B}$ telle que 100 % des spins soient déviés dans la direction de $\vec{B}$, et dans celui du photon qu'il n'existe pas de polariseur (λ, μ) laissant passer tous les photons avec une probabilité unité.

11.1.4 Non unicité de la préparation

Nous avons donné une préparation possible de la forme diagonale (11.5) de l'opérateur statistique : chaque état $|i\rangle$ apparaît avec la probabilité p_i. Mais, dès que les valeurs propres sont dégénérées, nous voyons que cette préparation ne peut pas être unique : si une valeur propre p_k est dégénérée, nous pouvons choisir comme états des superpositions linéaires arbitraires dans le sous-espace de cette valeur propre. En fait, la situation est encore bien plus ambiguë, car rien n'oblige à choisir des états orthogonaux : on peut parfaitement utiliser la décomposition (11.2) avec $\langle\varphi_\alpha|\varphi_\beta\rangle \neq \delta_{\alpha\beta}$. Une préparation détermine ρ, mais à un opérateur statistique donné correspondent une infinité de préparations possibles. Donnons un exemple simple sur le cas du spin 1/2. Supposons qu'un expérimentateur, Bob, envoie à sa collègue Alice une série de spins 1/2 qui ont une probabilité p de se trouver dans l'état $|+\rangle$, et une probabilité $\mathsf{q} = 1 - \mathsf{p}$ de se trouver dans l'état $|-\rangle$. On peut, par exemple, imaginer que Bob utilise un filtre de Stern-Gerlach pour préparer des spins dans l'état $|+\rangle$, et qu'il dispose

à la sortie du filtre d'un champ magnétique oscillant qui soumet les spins à une impulsion π de façon aléatoire, telle que l'oscillation de Rabi soit effective avec une probabilité q, et qu'elle ne soit pas apppliquée avec une probabilité p. Tous les résultats des mesures que peut faire Alice s'expriment en fonction de la matrice statistique

$$\rho = \mathsf{p}|+\rangle\langle+| + \mathsf{q}|-\rangle\langle-| = \begin{pmatrix} \mathsf{p} & 0 \\ 0 & \mathsf{q} \end{pmatrix} \tag{11.26}$$

En résumé, la préparation est la suivante : le spin apparaît dans l'état $|+\rangle$ avec la probabilité p et dans l'état $|-\rangle$ avec la probabilité $\mathsf{q} = 1 - \mathsf{p}$. Afin d'exhiber une autre préparation possible, définissons les états $|R\rangle$ et $|L\rangle$, $\langle L|R\rangle \neq 0$

$$|R\rangle = \sqrt{\mathsf{p}}\ |+\rangle + \sqrt{\mathsf{q}}\ |-\rangle \qquad |L\rangle = \sqrt{\mathsf{p}}\ |+\rangle - \sqrt{\mathsf{q}}\ |-\rangle \tag{11.27}$$

Les projecteurs $|R\rangle\langle R|$ et $|L\rangle\langle L|$ sont

$$|R\rangle\langle R| = \begin{pmatrix} \mathsf{p} & \sqrt{\mathsf{pq}} \\ \sqrt{\mathsf{pq}} & \mathsf{q} \end{pmatrix} \qquad |L\rangle\langle L| = \begin{pmatrix} \mathsf{p} & -\sqrt{\mathsf{pq}} \\ -\sqrt{\mathsf{pq}} & \mathsf{q} \end{pmatrix}$$

et manifestement, si Bob envoie à Alice des spins dans l'état $|R\rangle$ ou dans l'état $|L\rangle$ avec une probabilité de 50 %, l'opérateur ρ est inchangé

$$\rho = \frac{1}{2}\left(|R\rangle\langle R| + |L\rangle\langle L|\right) \tag{11.28}$$

Dans cette autre préparation correspondant au même opérateur statistique, le spin apparaît dans l'état $|R\rangle$ avec la probabilité 1/2 et dans l'état $|L\rangle$ avec la probabilité 1/2. Alice n'a aucun moyen de déterminer quelle est la préparation utilisée par Bob. Un autre exemple de la non unicité des préparations est donné dans l'exercice 11.6.5. Si l'on reprend l'exemple de la distribution canonique (11.4), il est habituel de dire que "le système se trouve dans l'état E_r avec une probabilité $\mathsf{p}(E_r) = Z^{-1}\exp(E_r/k_{\mathrm{B}}T)$", mais cette affirmation, bien qu'intuitivement commode, ne correspond qu'à une préparation particulière, et il existe une infinité d'autres préparations donnant le même opérateur statistique.

En outre, il existe une autre situation, entièrement différente, où l'état de spin dont dispose Alice est décrit par la matrice statistique ρ (11.26). Supposons en effet qu'Alice et Bob disposent de paires de spins dans un état intriqué $|\Phi_{AB}\rangle$; l'opérateur statistique d'Alice est

$$\rho(\equiv \rho_A) = \mathrm{Tr}_{\mathcal{B}}|\Phi_{AB}\rangle\langle\Phi_{AB}| \tag{11.29}$$

Alice effectue des mesures sur le spin $\mathcal{A}$ et Bob sur le spin $\mathcal{B}$. Choisissons pour $|\Phi_{AB}\rangle$

$$|\Phi_{AB}\rangle = \sqrt{\mathsf{p}}\ |+_A \otimes +_B\rangle + \sqrt{\mathsf{q}}|-_A \otimes -_B\rangle \tag{11.30}$$

L'opérateur statistique (11.47) coïncide avec (11.44)

$$\mathrm{Tr}_{\mathcal{B}}|\Phi_{AB}\rangle\langle\Phi_{AB}| = \mathsf{p}|+_A\rangle\langle+_A| + \mathsf{q}|-_A\rangle\langle-_A|$$

Si Alice est limitée à des mesures sur le spin $\mathcal{A}$, elle n'a aucun moyen de distinguer entre la situation où Bob lui envoie des spins orientés de façon aléatoire, et celle où elle mesure le spin $\mathcal{A}$ d'une paire intriquée $\mathcal{AB}$ dans l'état (11.30). Le vecteur $|\Phi_{AB}\rangle$ est appelé une *purification* de l'opérateur statistique ρ_A. Il existe une infinité de façons d'écrire la purification, par exemple, en utilisant la base $\{|+,\hat{x}_B\rangle, |-,\hat{x}_B\rangle\}$ qui diagonalise σ_{xB}. Un calcul immédiat permet de mettre $|\Phi_{AB}\rangle$ sous la forme

$$|\Phi_{AB}\rangle = \frac{1}{\sqrt{2}}\left(|R_A \otimes +,\hat{x}_B\rangle + |L_A \otimes -,\hat{x}_B\rangle\right) \qquad (11.31)$$

Supposons que Bob mesure la composante σ_{zB} du spin $\mathcal{B}$. S'il trouve la valeur $+1$ (avec une probabilité p), alors Alice recevra le spin $\mathcal{A}$ dans l'état $|+_A\rangle$, et s'il mesure la valeur -1 (avec une probabilité q), alors Alice recevra le spin $\mathcal{A}$ dans l'état $|-_A\rangle$. Ceci correspond à la préparation (11.26) de ρ. Si, au contraire, Bob mesure dans la base de σ_{xB}, Alice recevra le spin $\mathcal{A}$ dans l'état $|R_A\rangle$ ou $|L_A\rangle$ avec une probabilité de 50 %, ce qui correspond à la préparation (11.28) de ρ.

La généralisation de ces résultats a été énoncée par Gisin, Hughston, Josza et Wootters (GHJW) sous la forme du théorème suivant énoncé sans démonstration.

Théorème GHJW. Soit un opérateur statistique ρ agissant dans $\mathcal{H}_A$ de dimension d_A et N préparations

$$\rho = \sum_{i=1}^{D_n} \mathsf{p}_i^{(n)}|\varphi_i^{(n)}\rangle\langle\varphi_i^{(n)}| \qquad (11.32)$$

Il existe alors dans $\mathcal{H}_{AB}$ de dimension $\geq d_A \times D$, avec $D = \mathrm{Max}[D_n]$, un vecteur $|\Phi_{AB}\rangle$ et D_n propriétés physiques $B_i^{(n)}$ telles qu'une mesure de $B_i^{(n)}$ laisse le système $\mathcal{A}$ dans l'état $|\varphi_i^{(n)}\rangle$ avec une probabilité $\mathsf{p}_i^{(n)}$.

En résumé, lorsqu'Alice est limitée à des mesures sur le spin $\mathcal{A}$, elle peut déterminer ρ en effectuant un nombre suffisant de mesures, mais elle ne peut pas savoir si Bob lui envoie la préparation (11.26) de ρ ou la préparation (11.27) (ou une autre). Elle ne peut pas non plus savoir si Bob lui envoie, par exemple, une série spins qui se trouvent avec une probabilité $\mathsf{p}(\mathsf{q})$ dans l'état $|+\rangle$ ($|-\rangle$), ou si elle mesure le spin $\mathcal{A}$ de paires intriquées $\mathcal{AB}$ de deux spins dans l'état $|\Phi_{AB}\rangle$ (11.30). Cependant, les deux situations diffèrent radicalement. Dans le premier cas, chacun des spins $\mathcal{A}$ se trouve dans un état bien défini, même s'il est inconnu d'Alice : les opérateurs statistiques construits à partir de (11.2) seront appelés *mélanges propres*. Dans le second cas, l'état de spin n'est pas seulement inconnu d'Alice, *il n'existe pas* tant que Bob (ou Alice) n'ont effectué aucune mesure : dans une paire intriquée de spins 1/2, comme celle

décrite par (2.63), *l'état de spin d'une particule individuelle n'existe pas !* Un opérateur statistique obtenu en prenant une trace partielle sur un état pur de $\mathcal{H}_{AB}$ sera appelé *mélange impropre*. Cette distinction entre les deux types de mélange sera fondamentale pour la discussion de la mesure.

11.1.5 Dépendance temporelle de l'opérateur statistique

Il n'est pas difficile de trouver la loi d'évolution temporelle de l'opérateur statistique pour un système fermé. En effet, si nous considérons d'abord l'opérateur statistique d'un cas pur, nous obtenons l'équation d'évolution suivante pour le projecteur $\mathcal{P}_\varphi(t)$ en utilisant (4.11)

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_{\varphi(t)} = \mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}\Big(|\varphi(t)\rangle\langle\varphi(t)|\Big) = H(t)\mathcal{P}_{\varphi(t)} - \mathcal{P}_{\varphi(t)}H(t) = [H(t), \mathcal{P}_{\varphi(t)}]$$

En sommant sur les probabilités p_α on obtient l'équation d'évolution de $\rho(t)$

$$\boxed{\mathrm{i}\hbar\,\frac{\mathrm{d}\rho(t)}{\mathrm{d}t} = [H(t), \rho(t)]} \tag{11.33}$$

On obtient une loi équivalente en utilisant l'opérateur d'évolution $U(t,0)$ (4.14)

$$\rho(t) = U(t,0)\,\rho(t=0)\,U^{-1}(t,0) \tag{11.34}$$

Une telle évolution temporelle de l'opérateur statistique est appelée *évolution hamiltonienne*, ou *évolution unitaire* (*cf.* § 4.2.1). Il est intéressant d'observer qu'un état de désordre maximal est un invariant dynamique car $[H, \rho] = 0$ si $\rho = I/d$.

Comme exemple important, écrivons la loi d'évolution de l'opérateur statistique d'un spin 1/2 dans un champ magnétique constant. Avec $\vec{B}$ parallèle à Oz, le hamiltonien (3.62) s'écrit

$$H = -\frac{1}{2}\gamma\sigma_z$$

et l'équation d'évolution (11.33) devient, en utilisant les relations de commutation (3.52) et le vecteur de Bloch $\vec{b}$ (11.22)

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}[H, \rho] = -\frac{1}{2}\gamma B(b_x\sigma_y - b_y\sigma_x) \tag{11.35}$$

ce qui est équivalent à

$$\frac{\mathrm{d}b_x}{\mathrm{d}t} = -\gamma B b_y \qquad \frac{\mathrm{d}b_y}{\mathrm{d}t} = \gamma B b_x \qquad \frac{\mathrm{d}b_z}{\mathrm{d}t} = 0$$

soit sous forme vectorielle

$$\frac{\mathrm{d}\vec{b}}{\mathrm{d}t} = -\gamma\vec{B}\times\vec{b} \tag{11.36}$$

ce qui est exactement l'équation différentielle classique (3.31) donnant la précession de Larmor. Le vecteur de Bloch suit le même mouvement qu'un spin classique.

Dans notre discussion de la RMN au § 5.2.2, nous avons examiné un spin isolé. En fait, les spins se trouvent dans un environnement fluctuant à la température T et, en l'absence de champ de radiofréquences, ils sont décrits par un opérateur statistique ρ correspondant à l'équilibre thermique dans un champ magnétique constant $\vec{B}_0$

$$\rho \simeq \frac{1}{2}\left(I + \frac{\hbar\omega_0}{2k_{\rm B}T}\,\sigma_z\right) = \frac{1}{2}\left(I + \frac{1}{2}\,\delta\mathsf{p}\,\sigma_z\right) \tag{11.37}$$

où $\delta\mathsf{p}$ est la différence de populations $\delta\mathsf{p} = \mathsf{p}_+ - \mathsf{p}_-$ (5.45) entre les niveaux $|+\rangle$ et $|-\rangle$. Le vecteur de Bloch a pour composantes $\vec{b} = (0, 0, \delta\mathsf{p}/2)$. L'application d'une impulsion de radiofréquences résonante pendant un temps $t = \theta/\omega_1$ transforme ρ en ρ_θ

$$\rho \to \rho_\theta = U[\mathcal{R}_x(-\theta)]\,\rho U^\dagger\,[\mathcal{R}_x(-\theta)]$$

en raison de (5.35). Il est facile d'effectuer explicitement le produit de matrices, mais il est plus élégant d'utiliser (2.70)

$$\begin{aligned}
\mathrm{e}^{\,\mathrm{i}\theta\sigma_x/2}\sigma_z\mathrm{e}^{-\mathrm{i}\theta\sigma_x/2} &= \sigma_z + \frac{\mathrm{i}\theta}{2}[\sigma_x,\sigma_z] + \frac{1}{2!}\left(\frac{\mathrm{i}\theta}{2}\right)^2[\sigma_x,[\sigma_x,\sigma_z]] + \cdots \\
&= \sigma_z + \frac{\mathrm{i}\theta}{2}\,\sigma_y - \frac{1}{2!}\,\theta^2\sigma_z + \cdots = \sigma_z\,\cos\theta + \sigma_y\,\sin\theta
\end{aligned}$$

ce qui n'est pas autre chose que la loi de transformation des composantes y et z d'un vecteur dans une rotation d'angle $-\theta$ autour de Ox. On en déduit

$$\rho_\theta = \frac{1}{2}\left[I + \frac{1}{2}\,\delta\mathsf{p}(\sigma_z\,\cos\theta + \sigma_y\,\sin\theta)\right] \tag{11.38}$$

Dans le cas particulier d'une impulsion $\pi/2$ ($\theta = \pi/2$), le résultat est

$$\rho_{\pi/2} = \frac{1}{2}\left[I + \frac{1}{2}\delta\mathsf{p}\,\sigma_y\right] \tag{11.39}$$

Comme la matrice σ_y est non diagonale, on a créé des cohérences : *la différence de populations initiale a été convertie en cohérences*. Le retour à l'équilibre est contrôlé par le temps de relaxation T_2. Dans le cas d'une impulsion π, on obtient une inversion des populations des niveaux $|+\rangle$ et $|-\rangle$ et le retour à l'équilibre est contrôlé par le temps de relaxation T_1. On utilise de préférence une impulsion $\pi/2$ car en général $T_2 \ll T_1$.

11.1.6 Postulats

L'introduction de l'opérateur statistique permet de donner une formulation plus générale des postulats du chapitre 4.

- **Postulat Ia.** L'état d'un système quantique est représenté mathématiquement par un opérateur statistique ρ agissant dans un espace de Hilbert des états $\mathcal{H}$.
- **Postulat IIa (Règle de Born).** La probabilité p_χ de trouver le système quantique dans l'état $|\chi\rangle$ est donnée par

$$\mathsf{p}_\chi = \mathrm{Tr}\left(\rho|\chi\rangle\langle\chi|\right) = \mathrm{Tr}\left(\rho\,\mathcal{P}_\chi\right) \tag{11.40}$$

- **Postulat IVa.** L'évolution temporelle de l'opérateur statistique est donnée par (11.33), le postulat **III** restant inchangé.

Soulignons à nouveau que (11.33) ne vaut que pour un système fermé. L'évolution temporelle de l'opérateur statistique d'un système qui n'est pas fermé est beaucoup plus complexe et sera examinée au chapitre 18. En mécanique statistique, le cas d'un système en contact avec un réservoir de chaleur donne un exemple typique de système qui n'est pas fermé. L'évolution de l'ensemble système+réservoir est unitaire (si cet ensemble est fermé !), mais celle du système, obtenue en prenant la trace sur les variables du réservoir, ne l'est pas[12].

11.2 Inégalités de Bell

11.2.1 Démonstration de l'inégalité BCHSH

La forme de la théorie quantique telle que nous la connaissons aujourd'hui fut établie dans les années 1925-1927 dans le cas particulier de la mécanique ondulatoire, principalement par Heisenberg, Schrödinger et Born, mais ce fut Bohr qui en fit la synthèse en donnant une interprétation épistémologique à tous les développements techniques accumulés durant ces trois années. Cette interprétation est implicite dans la présentation des postulats au chapitre 4 ; nous en donnerons une évaluation critique ultérieurement (§ 11.4.7). Bien qu'il reconnût l'efficacité de la mécanique quantique et ses succès, par exemple en physique atomique et moléculaire, Einstein était en profond désaccord avec l'interprétation de Bohr. Pendant plusieurs années, il s'efforça de trouver des insuffisances dans la formulation de la théorie quantique telle qu'elle avait été plus ou moins acceptée par la grande majorité des physiciens, mais ses objections furent balayées par Bohr. Cependant, dans un article publié en 1935 avec Podolsky et Rosen, article universellement connu sous l'acronyme EPR formé par les initiales des trois auteurs, Einstein pensa avoir détecté

12. Dans une évolution hamiltonienne, l'entropie de von Neumann $-\mathrm{Tr}\,\rho\ln\rho$ est conservée, mais ce n'est pas le cas pour une évolution non hamiltonienne : l'entropie de von Neumann d'un système en contact avec un réservoir de chaleur n'est pas constante.

une faille : sans contester les succès de la mécanique quantique, les auteurs de l'article EPR estimaient que celle-ci ne pouvait pas être considérée comme une description complète de la réalité physique, suivant un argument qui sera développé au § 11.2.4. La réponse obscure et pour partie hors sujet de Bohr montra qu'Einstein avait visé juste. Cependant, pendant près de 30 ans, le débat resta uniquement sur le terrain philosophique et fut d'ailleurs oublié de la plupart des physiciens comme n'ayant aucune implication pratique. C'est seulement en 1964 que le physicien irlandais John Bell se rendit compte que les hypothèses contenues dans l'article EPR pouvaient être testées expérimentalement. Ces hypothèses conduisaient en effet à des inégalités, les inégalités de Bell, contredites par les prédictions de la théorie quantique. Dès la fin des années 1970, plusieurs groupes expérimentaux s'attaquèrent au problème, et on sait aujourd'hui que, selon toute probabilité, les inégalités de Bell ne sont pas compatibles avec l'expérience : les hypothèses implicites dans l'article EPR sont infirmées par les résultats expérimentaux. Dans cette section, nous expliquerons à l'aide d'un exemple intuitif l'origine des inégalités de Bell et nous les comparerons aux prédictions de la théorie quantique et aux données expérimentales récentes. Le débat entre Einstein et Bohr va bien au-delà d'une simple controverse, somme toute aujourd'hui réglée, car il pose des problèmes de fond sur la description spatio-temporelle des processus quantiques[13]. De plus, les concepts introduits dans ce débat, et en particulier le concept d'intrication, sont au cœur d'une série de développements extrêmement importants de la physique quantique moderne, en particulier de tous ceux liés à l'information quantique (section 11.5).

La question de la *localité* est au cœur de l'article EPR et de l'argumentation de Bell. Cette notion ne peut être comprise rigoureusement que dans le cadre de la relativité restreinte, mais nous allons d'abord utiliser une approche intuitive simplifiée et nous reviendrons ultérieurement sur la description relativiste. Nous allons définir la localité de la façon suivante : si deux stations expérimentales sont suffisamment éloignées l'une de l'autre, les choix effectués par un expérimentateur de la première station ne peuvent avoir aucune influence sur les résultats obtenus dans la seconde station, pourvu qu'aucun signal n'ait le temps de se propager entre les deux stations. Mais cela ne veut pas dire que les résultats de deux stations ne peuvent pas exhiber des corrélations ! Donnons un exemple élémentaire : une boîte contient deux T-shirts, l'un rouge et l'autre vert, chacun enveloppé dans un sac opaque. Réunis à un certain instant en un point d'espace-temps S (S pour source), Alice et Bob prennent chacun dans la boîte un des deux T-shirts et, sans ouvrir leur sac, partent dans deux directions opposées et parcourent chacun 1.5 milliards de kilomètres. Ayant terminé leur périple, ils enfilent leur T-shirt. Si Alice constate qu'elle est habillée en rouge, elle sait que Bob est habillé en vert, et *vice versa*. Bien qu'un échange de signaux entre Alice et Bob prenne plus

13. Bohr a certes "gagné" le débat, ses conceptions ont prévalu sur celles d'Einstein, mais il n'a jamais compris la profondeur des idées contenues dans l'article EPR.

de 3 heures, Alice connaît instantanément la couleur du T-shirt de Bob. Bien évidemment, il n'y a aucun mystère : une corrélation entre les couleurs des T-shirts a été introduite à la source, il n'y a pas eu de propagation d'un signal entre les deux personnages et la corrélation observée est en fait la conséquence d'une information préétablie.

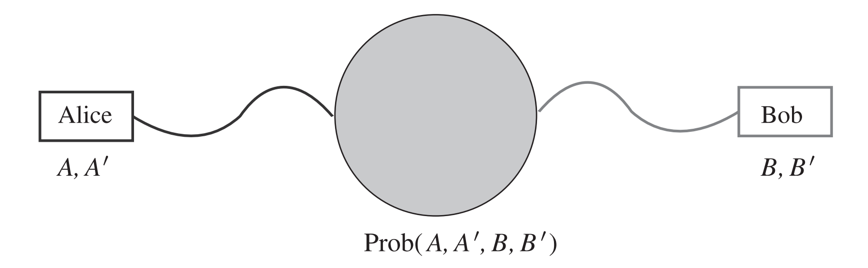


Fig. 11.2 – Stratégie de communication préétablie probabiliste.

Nous allons étoffer la garde-robe d'Alice et de Bob en leur attribuant non seulement des T-shirts mais également des pantalons. Ces T-shirts et ces pantalons sont comme ci-dessus soit verts, soit rouges, A et A' désignent les couleurs du T-shirt et du pantalon d'Alice, B et B' celles du T-shirt et du pantalon de Bob. Il est commode d'attribuer la valeur $+1$ ou -1 à chacune des couleurs : par exemple $A = +1$ lorsque le T-shirt d'Alice est rouge, $B' = -1$ lorsque le pantalon de Bob est vert. Le processus est schématisé sur la figure 11.2 : les habits sont distribués selon une loi de de probabilité $\mathsf{p}(A, A', B, B')$ qui constitue l'information préétablie, et dont la conséquence est l'existence de corrélations entre les habits d'Alice et de Bob. Comme ci-dessus, les habits sont enveloppés dans des sacs opaques et Alice et Bob les enfilent quand ils sont suffisamment éloignés pour qu'ils ne puissent échanger aucune information au moment où ils s'habillent. Il sera commode de supposer la propriété de symétrie vérifiée par tous les dispositifs étudiés par la suite

$$\mathsf{p}(A) = \mathsf{p}(-A) \qquad \mathsf{p}(A) = \sum_{A', B, B'} \mathsf{p}(A, A', B, B')$$

et une équation analogue pour les 3 autres variables. L'existence d'une loi de probabilité permet de montrer une inégalité remarquable, l'inégalité BCHSH (Bell, Clauser, Horne, Shimony et Holt) en partant de l'égalité triviale

$$X = A(B + B') + A'(B - B') = \pm 2 \tag{11.41}$$

En effet, soit $B = B'$ auquel cas $X = \pm 2A = \pm 2$, soit $B = -B'$, auquel cas $X = \pm 2A' = \pm 2$. La valeur moyenne de X, calculée avec la loi de probabilité $\mathsf{p}(A, A', B, B')$ est nécessairement inférieure ou égale à 2 en module

$$S = |\langle X \rangle| = |\langle AB \rangle + \langle AB' \rangle + \langle A'B \rangle - \langle A'B' \rangle| \leq 2 \tag{11.42}$$

Comme $\langle A \rangle = \cdots = \langle B' \rangle = 0$, $S = 0$ correspond à l'absence de corrélations entre les résultats d'Alice et de Bob. Dans le schéma de la figure 11.2, les corrélations proviennent d'une stratégie de communication partagée probabiliste, que l'on appelle parfois un aléatoire partagé : les fluctuations des résultats proviennent des fluctuations des sources.

11.2.2 Physique quantique et borne de Cirelson

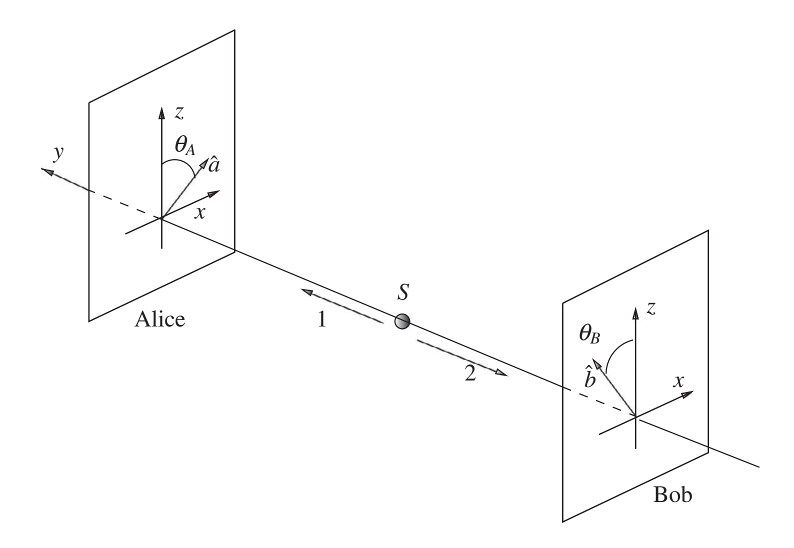


FIG. 11.3 – Configuration des axes pour l'expérience de mesure de deux spins. La source S émet une paire de particules de spin 1/2 intriqués, et ces spins sont mesurés suivant les directions $\hat{a}\,(\theta_A)$ et $\hat{b}\,(\theta_B)$.

Nous allons comparer ce résultat (11.42) à celui obtenu dans un exemple de physique quantique qui, au premier abord, semble présenter une analogie parfaite avec la situation décrite au § 11.2.1. Considérons l'état intriqué $|\Phi\rangle$ (2.63) de deux spins 1/2 $\mathcal{A}$ et $\mathcal{B}$ que nous récrivons ci-dessous

$$|\Phi\rangle = \frac{1}{\sqrt{2}}(|+\otimes-\rangle - |-\otimes+\rangle) \tag{11.43}$$

et supposons qu'à l'aide d'analyseurs de Stern-Gerlach, Alice et Bob mesurent respectivement la composante du spin $\mathcal{A}$ et celle du spin $\mathcal{B}$ suivant des directions $\hat{a}$ et $\hat{b}$ dans le plan (x, z). La forme de $|\Phi\rangle$ implique une anticorrélation parfaite entre les composantes Oz des deux spins, et compte tenu de l'invariance par rotation de $|\Phi\rangle$ (exercice 2.5.14), la probabilité jointe de trouver les spins orientés dans la même direction ($\hat{a} \equiv \hat{b}$) est nulle quelle que soit cette direction. Si le spin d'Alice est up, celui de Bob est down et *vice versa*, ce qui rappelle le premier exemple des T-shirts : si celui d'Alice est vert, celui de

Bob est rouge et *vice versa*. Afin de simplifier la géométrie, nous supposons que $\hat{a}$ et $\hat{b}$ sont situés dans le plan xOz et font un angle θ_A et θ_B avec l'axe Oz (figure 11.3). Nous allons calculer la probabilité jointe p_{++} que le spin d'Alice soit up pour la direction $\hat{a}$ et celui de Bob également up pour la direction $\hat{b}$. Nous savons déjà que $\mathsf{p}_{++} = 0$ si $\hat{a} \parallel \hat{b}$, et nous nous intéressons au cas où $\hat{a}$ et $\hat{b}$ ont des orientations arbitraires. Nous pouvons utiliser l'invariance par rotation de $|\Phi\rangle$ pour choisir $\hat{a}$ parallèle à Oz ($\theta_A = 0$), tandis que l'angle entre Oz et $\hat{b}$ est égal à $\theta_B = \theta$. D'après (3.59), l'état $|+,\theta\rangle$ obtenu à partir de $|+\rangle$ par une rotation de θ autour de Oy est

$$|+,\theta\rangle = \cos\frac{\theta}{2}\,|+\rangle + \sin\frac{\theta}{2}\,|-\rangle \tag{11.44}$$

et le produit tensoriel $|+\otimes+,\theta\rangle$ vaut

$$|+\otimes+,\theta\rangle = \cos\frac{\theta}{2}\,|+\otimes+\rangle + \sin\frac{\theta}{2}\,|+\otimes-\rangle \tag{11.45}$$

L'amplitude de probabilité $a(|\Phi\rangle \to |+\otimes+,\theta\rangle)$ est donnée par le produit scalaire

$$\langle +\otimes+,\theta|\Phi\rangle = \frac{1}{\sqrt{2}}\sin\frac{\theta}{2}$$

d'où la probabilité

$$\mathsf{p}_{++} = \mathsf{p}(|\Phi\rangle \to |+\otimes+,\theta\rangle) = \frac{1}{2}\sin^2\frac{\theta}{2}$$

Pour $\theta = 0$, nous retrouvons bien $\mathsf{p}_{++} = 0$. Par symétrie, nous déduisons de cette équation

$$\mathsf{p}_{++} = \mathsf{p}_{--} = \frac{1}{2}\sin^2\frac{\theta}{2} \qquad \mathsf{p}_{+-} = \mathsf{p}_{-+} = \frac{1}{2}\cos^2\frac{\theta}{2} \tag{11.46}$$

Si nous notons comme précédemment $A = \pm 1$ et $B = \pm 1$, les résultats des mesures d'Alice et Bob ($A, B = +1$ pour un spin up, $A, B = -1$ pour un spin down) la valeur moyenne $\langle AB\rangle$ de AB se déduit de (11.46)

$$\langle AB\rangle = \mathsf{p}_{++} + \mathsf{p}_{--} - \mathsf{p}_{+-} - \mathsf{p}_{-+} = -\cos\theta = -\hat{a}\cdot\hat{b} \tag{11.47}$$

Un autre façon d'écrire ce résultat est de remarquer que A et B sont les deux valeurs propres possibles des opérateurs $\vec{\sigma}\cdot\hat{a}$ et $\vec{\sigma}\cdot\hat{b}$ et que, par conséquent,

$$\langle AB\rangle = \langle\Phi|(\vec{\sigma}\cdot\hat{a})\otimes(\vec{\sigma}\cdot\hat{b})|\Phi\rangle = \langle(\vec{\sigma}\cdot\hat{a})\otimes(\vec{\sigma}\cdot\hat{b})\rangle = -\hat{a}\cdot\hat{b} \tag{11.48}$$

Le résultat (11.48) est donc la valeur moyenne dans l'état $|\Phi\rangle$ de l'opérateur à deux particules $(\vec{\sigma}\cdot\hat{a})\otimes(\vec{\sigma}\cdot\hat{b})$, ce qui est dans la droite ligne du point de vue de Bohr selon lequel *on doit considérer l'ensemble du dispositif expérimental*, et donc l'ensemble formé par les deux détecteurs éventuellement très éloignés

qui effectuent la mesure de cet opérateur. Les résultats d'Alice et de Bob sont ensuite rassemblés par une procédure classique, où l'information se propage à une vitesse inférieure à c (par exemple une ligne téléphonique), et on peut alors examiner les corrélations éventuelles entre les résultats des deux détecteurs. Une méthode plus expéditive pour arriver au résultat (11.46) consiste à utiliser la réduction du paquet d'ondes (4.7). Supposons qu'Alice mesure un spin up suivant Oz : dans ces conditions, le vecteur $|\Phi\rangle$ (11.43) est transformé en $|+\otimes-\rangle$ avec une probabilité 1/2

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\left(|+\otimes-\rangle - |-\otimes+\rangle\right) \Longrightarrow |+\otimes-\rangle \tag{11.49}$$

En effet, si Alice a mesuré un spin up, la partie $|+\otimes-\rangle$ de $|\Phi\rangle$ est sélectionnée par la réduction du paquet d'ondes, ce qui entraîne que le spin $\mathcal{B}$ est dans l'état $|-\rangle$. Il suffit ensuite d'utiliser (11.44) pour retrouver (11.46)

$$\mathsf{p}_{++} = |\langle +,\theta|-\rangle|^2 = \frac{1}{2}\sin^2\frac{\theta}{2} \qquad \mathsf{p}_{+-} = |\langle -,\theta|-\rangle|^2 = \frac{1}{2}\cos^2\frac{\theta}{2}$$

Cependant, l'interprétation physique de cette réduction du paquet d'ondes pose un problème de fond : comment se fait-il que la mesure du spin d'Alice se traduise *instantanément* par une action sur le vecteur d'état du spin de Bob ? Avant la mesure d'Alice, le spin de Bob n'avait pas de valeur définie : il était décrit par l'opérateur statistique (11.19), après la mesure il se trouve dans un état bien défini $|-\rangle$ ou $|+\rangle$ selon le résultat d'Alice, même si Alice et Bob sont sur deux galaxies différentes. De plus, si les mesures d'Alice et Bob sont séparées par un intervalle du genre espace, l'ordre temporel des mesures n'est pas défini et dépend du référentiel d'inertie utilisé (figure 11.6). Il ne semble pas possible de considérer la réduction du paquet d'ondes comme un processus physique, à moins de violer la causalité : la mesure d'Alice ne peut pas influencer *l'état physique* du spin de Bob. La réduction du paquet d'ondes s'apparente plutôt à une acquisition d'information supplémentaire sur le vecteur d'état, superficiellement analogue à celle que l'on peut acquérir sur une distribution de probabilité classique, lorsque dans le cadre d'une interprétation bayesienne on actualise une probabilité afin de tenir compte d'une information nouvelle. On peut l'illustrer par l'exemple suivant, lorsque l'on connaît en mécanique classique la distribution de probabilité $\mathsf{p}(\vec{r})$ de la position d'une particule, et qu'ensuite cette particule est localisée en un point $\vec{r} = \vec{r}_0$: on passe instantanément d'une distribution étalée $\mathsf{p}(\vec{r})$ à une distribution $\delta(\vec{r} - \vec{r}_0)$, et la distribution de probabilité s'annnule instantanément pour $\vec{r} \neq \vec{r}_0$, ce qui est formellement un phénomène non local.

Supposons maintenant qu'Alice et Bob utilisent deux axes supplémentaires $\hat{a}'$ et $\hat{b}'$ pour leurs mesures : si Alice mesure dans la direction $\hat{a}$, elle obtient le résultat $A' = \pm 1$, et si Bob mesure dans la direction $\hat{b}'$, il obtient le résultat $B' = \pm 1$. Nous pouvons évaluer théoriquement la quantité S définie en (11.42), et cette quantité peut être déterminée expérimentalement par Alice

et Bob, en effectuant une série de mesures avec un choix d'axes $(\hat{a}, \hat{b})$, puis $(\hat{a}, \hat{b}')$, puis $(\hat{a}', \hat{b})$ et enfin $(\hat{a}', \hat{b}')$. Si nous notons A_n et B_n, les résultats de la mesure numéro n pour le choix d'axes $(\hat{a}, \hat{b})$ par exemple, la valeur moyenne $\langle AB\rangle$ est tirée de

$$\langle AB\rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} A_n B_n \tag{11.50}$$

et une équation analogue est valable pour les trois autres choix d'axes.

Nous allons voir qu'en physique quantique la quantité S définie en (11.42) peut atteindre la valeur $2\sqrt{2}$, et que, par conséquent, l'inégalité BCHSH (11.42) est violée. Cette valeur $2\sqrt{2}$ est la *borne de Cirelson* que l'on démontre en examinant la quantité

$$\begin{aligned} Q &= (\vec{\sigma}\cdot\hat{a}) \otimes (\vec{\sigma}\cdot\hat{b}) + (\vec{\sigma}\cdot\hat{a}) \otimes (\vec{\sigma}\cdot\hat{b}') + (\vec{\sigma}\cdot\hat{a}') \otimes (\vec{\sigma}\cdot\hat{b}) - (\vec{\sigma}\cdot\hat{a}') \otimes (\vec{\sigma}\cdot\hat{b}') \\ &= (\vec{\sigma}\cdot\hat{a}) \otimes (\vec{\sigma}\cdot\vec{B}) + (\vec{\sigma}\cdot\hat{a}') \otimes (\vec{\sigma}\cdot\vec{B}') \end{aligned} \tag{11.51}$$

avec $\vec{B} = \hat{b} + \hat{b}'$ et $\vec{B}' = \hat{b} - \hat{b}'$ qui vérifient

$$\vec{B}\cdot\vec{B}' = 0 \qquad \vec{B}\times\vec{B}' = -2\hat{b}\times\hat{b}'$$

On observe que le module $|\langle Q\rangle|$ de la valeur moyenne de Q n'est autre que la quantité S (11.42).

$$|\langle Q\rangle| = |\langle\Psi|Q|\Psi\rangle| = |\langle AB\rangle + \langle AB'\rangle + \langle A'B\rangle - \langle A'B'\rangle|$$

où $|\Psi\rangle$ est un état arbitraire des deux spins, qui pourrait même être un opérateur statistique ρ : $\langle Q\rangle = \mathrm{Tr}\,(\rho\, Q)$. Calculons Q^2 en utilisant

$$[(\vec{\sigma}\cdot\hat{a})(\vec{\sigma}\cdot\hat{a}')] \otimes [(\vec{\sigma}\cdot\vec{B})(\vec{\sigma}\cdot\vec{B}')] = [\hat{a}\cdot\hat{a}' + \mathrm{i}\vec{\sigma}\cdot(\hat{a}\times\hat{a}')] \otimes [\mathrm{i}\vec{\sigma}\cdot(\vec{B}\times\vec{B}')]$$

et deux identités analogues, avec comme résultat

$$\langle Q\rangle^2 \le \langle Q^2\rangle = 4 - 4\langle\Psi|[\vec{\sigma}\cdot(\hat{a}\times\hat{a}')] \otimes [\vec{\sigma}\cdot(\hat{b}\times\hat{b}')]|\Psi\rangle \le 8$$

Lorsque les deux spins sont dans l'état $|\Phi\rangle$ (11.43), on déduit de (11.48)

$$\begin{aligned} \langle Q\rangle^2 \le \langle Q^2\rangle &= 4[1 + (\hat{a}\times\hat{a}')\cdot(\hat{b}\times\hat{b}'] \\ &= 4[1 - (\hat{a}\cdot\hat{b})(\hat{a}'\cdot\hat{b}') + (\hat{a}'\cdot\hat{b})(\hat{a}\cdot\hat{b}'] \end{aligned} \tag{11.52}$$

La configuration de la figure 11.4 pour les orientations $(\hat{a}, \hat{a}', \hat{b}, \hat{b}')$, avec

$$(\hat{a}\cdot\hat{b}) = (\hat{a}'\cdot\hat{b}) = (\hat{a}\cdot\hat{b}') = -(\hat{a}'\cdot\hat{b}') = \frac{1}{\sqrt{2}}$$

permet d'atteindre la borne de Cirelson $|\langle Q\rangle| = \sqrt{8} = 2\sqrt{2}$.

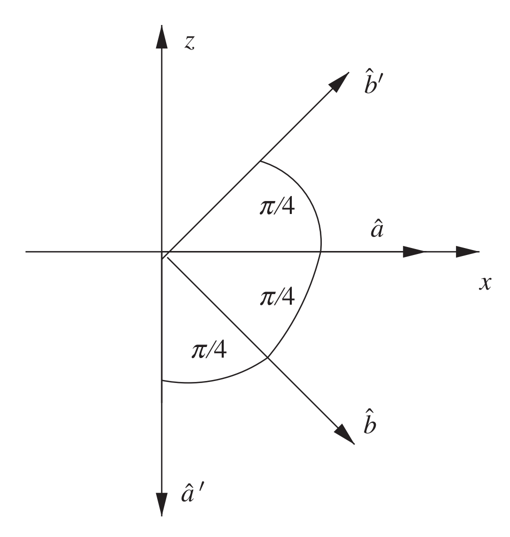


FIG. 11.4 – Configuration des axes permettant d'atteindre la borne de Cirelson.

La borne de Cirelson viole l'inégalité de Bell, alors que l'on semblait disposer d'une analogie parfaite entre l'exemple du § 11.2.1 et les spins intriqués : la valeur du spin suivant $\hat{a}(\hat{a}')$ correspond à la couleur $A = \pm 1$ du T-shirt d'Alice ($A' = \pm 1$ de son pantalon), et de même pour Bob, alors que l'inégalité (11.42) est violée par la physique quantique. La différence entre le cas de l'information préétablie et le cas quantique est subtile, mais essentielle : pour évaluer par exemple la valeur moyenne $\langle AB' \rangle$ en effectuant un grand nombre de mesures successives (voir (11.50)), Alice et Bob peuvent se contenter d'observer la couleur de leur T-shirt et pantalon respectifs et ignorer celles de leur pantalon (Alice) et de leur T-shirt (Bob), mais l'information sur ces couleurs existe, même si elle n'est pas utilisée. Dans l'expérience avec les spins 1/2 intriqués, lorsqu'Alice choisit de mesurer le spin suivant $\hat{a}$ (elle observe la couleur de son T-shirt), elle s'interdit de mesurer le spin suivant $\hat{a}'$ *pour la même paire de spins* : les mesures suivant $\hat{a}$ et $\hat{a}'$ sont incompatibles, on peut effectuer soit l'une, soit l'autre, mais pas les deux à la fois. Si nous voulions pousser jusqu'au bout la correspondance T-shirt/pantalon et choix des axes pour les mesures de spin, il nous faudrait admettre que la valeur du spin suivant $\hat{a}'$ existe, même si Alice ne peut pas l'observer avec le choix d'axe $\hat{a}$: cette hypothèse est appelée *contrafactualité*. Avec cette hypothèse, l'inégalité de Bell est valable. La physique quantique est incompatible avec la contrafactualité : le choix de l'orientation $\hat{a}$ *exclut* celui de l'orientation $\hat{a}'$, les résultats de la mesure du spin ne préexistent pas à cette mesure, et cela n'a aucun sens de prétendre qu'une mesure non effectuée pourrait avoir un résultat. La violation de l'inégalité de Bell par la physique quantique montre que la distribution de probabilité $\mathsf{p}(A, A', B, B')$ n'existe pas. On appelle sou-

vent *théorème de Bell* la violation de ses inégalités par la physique quantique. C'est l'utilisation d'orientations différentes pour $\hat{a}$ et $\hat{b}$ qui est cruciale pour le théorème de Bell : avec un modèle probabiliste classique, on reproduit sans difficulté les (anti-)corrélations pour toutes les orientations des axes pourvu que $\hat{a} \parallel \hat{b}$.

11.2.3 Expériences avec des photons

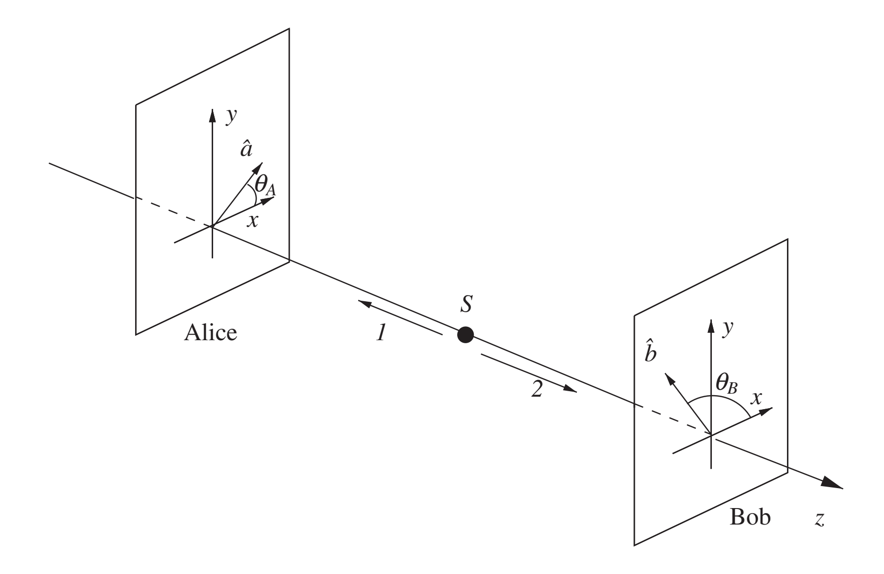


FIG. 11.5 – Définition des axes pour l'expérience avec des photons.

Il reste bien sûr à vérifier que l'expérience se conforme aux résultats de la théorie quantique. Les spins 1/2 sont commodes pour une discussion théorique, mais en pratique les expériences se font le plus souvent avec des paires de photons intriqués en polarisation. Les premières expériences concluantes ont été réalisées avec des paires produites dans une cascade atomique (Freedman et Clauser [1972], Clauser [1976], Aspect *et al.* [1982]). On utilise aujourd'hui une technique différente, la conversion paramétrique dans un cristal non linéaire. Dans l'expérience de Weihs *et al.* [1998], la conversion paramétrique produit une paire de photons dans un état intriqué en polarisation (figure 11.5)

$$|\Psi_{AB}\rangle = \frac{1}{\sqrt{2}}\big(|x_A y_B\rangle - |y_A x_B\rangle\big) \tag{11.53}$$

où $|x\rangle$ et $|y\rangle$ sont des états de polarisation linéaire suivant Ox et Oy. Cet état est invariant par rotation autour de Oz. Soit en effet $|\theta\rangle$ et $|\theta_\perp\rangle$ les états de polarisation (3.23)-(3.24) pour un photon polarisé suivant les directions θ

et $\theta_\perp$

$$\begin{aligned} |\theta\rangle &= \cos\theta\,|x\rangle + \sin\theta\,|y\rangle \\ |\theta_\perp\rangle &= -\sin\theta\,|x\rangle + \cos\theta\,|y\rangle \end{aligned} \tag{11.54}$$

Un calcul immédiat montre que $|\Psi_{AB}\rangle$ est invariant par rotation autour de Oz

$$|\Psi_{AB}\rangle = \frac{1}{\sqrt{2}}\big(|\theta_A\theta_{\perp B}\rangle - |\theta_{\perp A}\theta_B\rangle\big) \tag{11.55}$$

La polarisation des photons est analysée par Alice et Bob qui utilisent des axes $\hat{a}$ et $\hat{b}$ orientés suivant les directions θ_A et θ_B du plan xOy (figure 11.5). Introduisons la quantité A définie comme suit : $A = +1$ si la polarisation du photon est orientée suivant θ_A et $A = -1$ si elle est orientée suivant $\theta_{\perp A}$; on définit de même pour Bob les valeurs $B = \pm 1$. Comme dans le cas du spin $1/2$, on utilise l'invariance par rotation autour de Oz pour choisir $\hat{a}$ suivant Ox et $\hat{b}$ dans xOy faisant un angle θ avec Ox. On doit calculer la probabilité

$$\mathsf{p}\big(|\Psi_{AB}\rangle \to |x_A \otimes \theta_B\rangle\big)$$

avec

$$|x_A \otimes \theta_B\rangle = \cos\theta|x_A \otimes x_B\rangle + \sin\theta|x_A \otimes y_B\rangle$$

soit pour l'amplitude de probabilité

$$a\big(|\Psi_{AB}\rangle \to |x_A \otimes \theta_B\rangle\big) = \frac{1}{\sqrt{2}}\sin\theta$$

et pour la probabilité

$$\mathsf{p}\big(|\Psi_{AB}\rangle \to |x_A \otimes \theta_B\rangle\big) = \frac{1}{2}\sin^2\theta \tag{11.56}$$

Comme on pouvait s'y attendre, il suffit de remplacer l'angle $\theta/2$ du résultat (11.46) pour le spin $1/2$ par l'angle θ pour le photon de spin 1 et l'on trouve

$$\langle AB\rangle = \frac{1}{2}\Big[-2\cos^2\theta + 2\sin^2\theta\Big] = -\cos(2\theta) \tag{11.57}$$

La borne de Cirelson $S \le 2\sqrt{2}$ est évidemment identique à celle obtenue dans le cas du spin $1/2$, et l'orientation des axes donnant cette borne s'obtient à partir de celle de la figure 11.4 en divisant les angles par deux.

Revenons maintenant à la question de la localité dans le cadre de la relativité restreinte. Il est bien connu que la relativité restreinte interdit tout échange d'énergie ou d'information à une vitesse supérieure à la vitesse de la lumière, c. Autrement dit, la relativité restreinte interdit toute *communication supraluminique* (ce qui est souvent appelé en anglais condition de “non

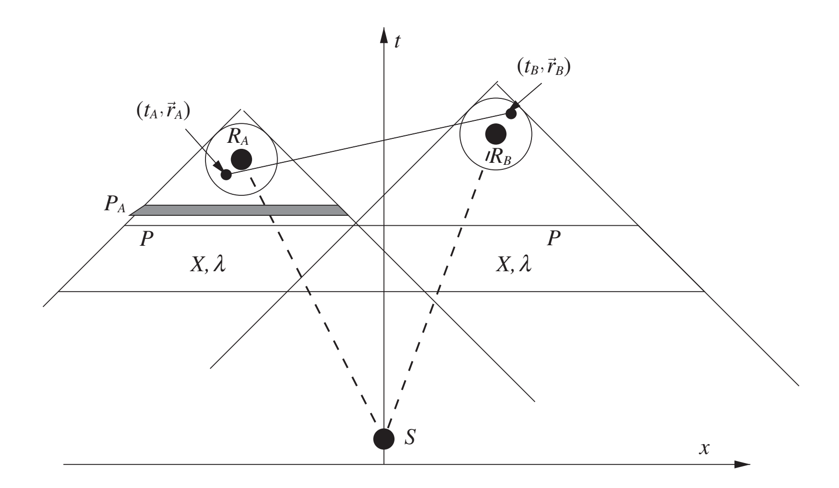


FIG. 11.6 – Espace-temps à deux dimensions (t, x). Les régions R_A et R_B sont séparées par un intervalle de genre espace : pour tout couple de points dans ces régions, $c^2(t_A - t_B)^2 - (\vec{r}_A - \vec{r}_B)^2 < 0$. La figure montre les cônes passés de R_A et R_B ; $t_B > t_A$, mais l'ordre peut être inversé dans un référentiel d'inertie différent. La source S émet une paire de particules dont la vitesse $v \leq c$.

signalling") : nous l'appellerons la condition de *causalité relativiste* [14] (ou simplement causalité). En relativité restreinte, un *événement* de coordonnés $(t, \vec{r})$ est défini par l'instant t et le point d'espace $\vec{r}$ où il se produit : par exemple, un accident de voiture est défini par l'heure et le lieu où il se produit. Deux événements peuvent être reliés de *façon causale* s'il est possible qu'une information se propage de l'un à l'autre à une vitesse inférieure ou égale à la vitesse de la lumière. Dans le cas contraire, il ne peut y avoir de relation de cause à effet entre les deux événements, c'est-à-dire qu'aucun des deux événements n'a pu influencer l'autre.

Nous allons donner une interprétation géométrique aux considérations précédentes en anticipant sur des notions qui seront développées dans la section 19.1. Soit deux événements, l'un O de coordonnées $(t = 0, \vec{r} = 0)$ et l'autre E de coordonnées $(t, \vec{r})$. Comme c est la vitesse limite de propagation de l'énergie ou de l'information, l'événement O ne peut influencer E que si $t > 0$ et $c^2t^2 - \vec{r}^2 \geq 0$. Si $c^2t^2 - \vec{r}^2 < 0$, on dit que les deux événements sont séparés par un *intervalle du genre espace* : E ne peut exercer aucune influence sur O et réciproquement. La surface $c^2t^2 - \vec{r}^2 = 0$ est appelée *cône de lumière* de O, l'intérieur de ce cône pour $t > 0$ $(t < 0)$ est le *cône futur (passé)* de O.

14. Il faut cependant bien comprendre que la relativité restreinte n'interdit pas des *déplacements* s'effectuant à des vitesses supérieures à c. Un exemple simple est donné par le déplacement de la tache de quelques km de diamètre d'un faisceau laser illuminant la Lune. Si le laser tourne avec une vitesse angulaire supérieure à environ 10 tours par seconde, alors la tache balaie la Lune avec une vitesse supérieure à c. Mais, il est impossible d'utiliser ce déplacement pour échanger de l'information, en dehors d'une information préétablie.

En résumé, O ne peut influencer que les événements situés dans son cône futur, et il ne peut avoir été influencé que par les événements de son cône passé. Toutes les théories physiques sont supposées obéir à ces propriétés. Par exemple, un champ électromagnétique classique dans la région d'espace-temps R_A de la figure 11.2 en l'absence de sources pour simplifier est entièrement déterminé par la donnée de ce champ dans la région P_A située dans le cône passé de R_A. Les théories qui vérifient ces propriétés sont appelées *théories locales*. Nous verrons, qu'en physique quantique, les résultats d'expériences dans des régions R_A et R_B séparées par un intervalle du genre espace ne sont pas déterminés uniquement par leurs cônes passés : *la physique quantique est non locale*[15].

Considérons deux expériences E_A et E_B réalisées dans ces deux régions d'espace-temps R_A et R_B : tout événement $x_A = (t_A, \vec{r}_A) \in R_A$ est séparé de tout événement $x_B = (t_B, \vec{r}_B) \in R_B$ par un intervalle de genre espace

$$c^2(t_A - t_B)^2 - (\vec{r}_A - \vec{r}_B)^2 < 0$$

Dans ces conditions, aucun choix fait par un expérimentateur de E_A ne peut influencer un résultat de E_B. Mais cela n'implique pas que les résultats de E_A et E_B ne peuvent pas être corrélés ! En effet, ces résultats peuvent avoir des causes communes situées dans l'intersection des cônes passés de R_A et R_B, par exemple l'événement S dans la figure 11.2 (S pour *source*)

Décrivons brièvement l'expérience de Weihs *et al.* [1998] (voir aussi Aspect [1999]) qui se rapproche le plus des conditions définies par la figure 11.6 pour tester les inégalités de Bell et le caractère non local de la physique quantique. Un photon ultraviolet est converti dans un cristal non linéaire en deux photons dans l'état de polarisation (11.53) : figure 11.7. Dans cette expérience, on peut modifier de façon aléatoire l'orientation des analyseurs pendant que les photons sont en vol entre leur point de production et les détecteurs, et aucune information sur l'orientation choisie par Alice ne peut être parvenue à Bob quand il effectue sa propre mesure, ce qui définit la condition de causalité. En effet, les deux détecteurs sont distants de 400 m, distance franchie par la lumière en 1.3 μs, alors que la durée des mesures ne dépasse pas 100 ns. Une expérience plus récente (Scheidl *et al.* [2008]), où les détecteurs sont distants de 144 km, est encore mieux placée pour vérifier la condition de causalité. Cependant, seulement 5 % des paires de photons sont détectées, et il faut admettre que ces 5 % constituent un échantillon représentatif : les conclusions

15. La terminologie est ambiguë et varie selon les auteurs. En théorie quantique des champs relativiste, on appelle théorie locale une théorie telle que deux opérateurs locaux, c'est-à-dire dépendant d'un point d'espace-temps $x = (t, \vec{r})$ et représentant des propriétés physiques, $O(x_A)$ et $O'(x_B)$, commutent si l'intervalle d'espace-temps $(x_A - x_B)$ est du genre espace, $[O(x_A), O'(x_B)] = 0$ (Streater et Wightman [1964]). Avec cette terminologie, la théorie quantique des champs relativiste et la mécanique quantique usuelle, qui en est la limite galiléenne, seraient des théories locales. C'est pourquoi, nous distinguons entre causalité relativiste qui implique la commutation des opérateurs locaux pour des intervalles du genre espace et la localité : la physique quantique obéit à la première propriété mais pas à la seconde.

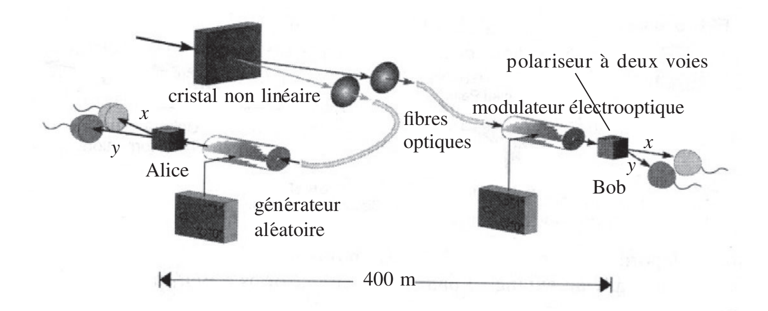


FIG. 11.7 – Expérience avec des photons intriqués. La paire de photons intriqués est produite dans un cristal non linéaire BBO et les deux photons partent dans des fibres optiques qui les amènent aux analyseurs de polarisation. Un modulateur électro-optique permet de choisir la direction d'analyse de la polarisation avant le passage dans les polariseurs. Deux générateurs de nombres aléatoires situés dans le voisinage d'Alice et de Bob choisissent au dernier moment l'orientation des axes d'analyse. D'après A. Zeilinger [1999].

de l'expérience peuvent être contestées en raison de cette "échappatoire de la détection". Si l'on voulait être absolument certain d'avoir prouvé la violation des inégalités de Bell, il faudrait éliminer dans une même expérience l'échappatoire de la détection et l'échappatoire de la causalité, c'est-à-dire détecter les paires de particules avec une probabilité d'au moins 80 % et réaliser les deux détections dans des régions séparées par un intervalle du genre espace. La figure 11.8 donne le schéma pour la condition de causalité, lorsque l'on tient compte de l'incertitude sur le temps de production de la paire : temps de cohérence du laser τ_c et de la durée τ_M de la mesure. Dans l'expérience de Weihs *et al.* [1998], la durée totale de la mesure inclut les contributions suivantes : durée de génération des nombres aléatoires, activation du modulateur électro-optique et déclenchement de l'avalanche des photo-électrons dans le détecteur. La somme de ces contributions est estimée à 75 ns, et l'incertitude due au temps de cohérence du laser, qui peut être réduite à 100 ps, est négligeable. Cependant, on pourrait exiger d'attendre jusqu'à la brisure de la chaîne de von Neumann (§ 11.4.4) pour estimer τ_M, et dans ce cas l'estimation deviendrait évidemment très délicate. Si l'on rejette cette dernière objection, on peut raisonnablement admettre que les expériences réalisées avec des photons vérifient la condition de causalité relativiste, mais elles ne vérifient pas celle de la détection ; au contraire, des expériences réalisées avec des ions ont une efficacité de détection proche de 100 %, mais les deux détecteurs sont distants de seulement un mètre (Matsukevitch *et al.* [2008]), ce qui ne permet pas de vérifier la condition de causalité.

À strictement parler, il est donc possible aujourd'hui de contester que l'expérience viole les inégalités de Bell. Toutefois, les mécanismes envisagés pour

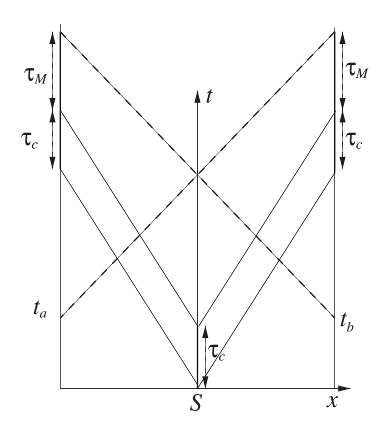


Fig. 11.8 – Condition de causalité. L'incertitude sur le temps de production de la paire est τ_c et le temps pris par la mesure est τ_M. Les tirets représentent la propagation d'un signal à une vitesse c. Le choix d'Alice pour l'orientation de son axe $\hat{a}$ doit être fait avant un temps t_a, de sorte que Bob ne peut recevoir aucune information sur ce choix avant d'avoir terminé sa mesure. D'après Scarani [2011].

éviter cette violation sont extrêmement artificiels et ne reposent sur aucune théorie plausible. La grande majorité des physiciens estime que cette violation est prouvée avec une grande probabilité. Cependant, il serait souhaitable de réaliser une expérience éliminant simultanément toutes les échappatoires, ce qui sera sans doute possible dans un délai de quelques années. Si l'on ne tient pas compte des échappatoires, on peut affirmer que *la théorie quantique est non locale, au sens où ses résultats ne peuvent être reproduits par aucune stratégie de communication préétablie*, mais cette non localité ne contredit jamais la causalité relativiste et ne permet pas de transmission d'information à une vitesse supérieure à celle de la lumière : Alice et Bob observent chacun une suite aléatoire de $+1$ et de -1, qui ne contient aucune information, et c'est seulement en comparant leurs résultats transmis par une voie classique, à une vitesse inférieure ou égale à c, qu'ils peuvent se rendre compte de leur corrélation. On trouvera des commentaires additionnels sur ce point dans les § 11.3.1 et § 11.5.1. En dernière analyse, et même si l'on évite de faire appel au mécanisme explicitement non local de réduction du paquet d'ondes, la non localité vient de ce que le dispositif expérimental possède une extension spatiale irréductible : la distance entre les deux détecteurs peut être arbitrairement grande.

11.2.4 EPR et la non localité quantique

Cette sous-section réexamine l'argument EPR vu comme conséquence de la localité en reprenant l'exemple des spins $1/2$ dans la configuration de la figure 11.6. Lorsqu'Alice mesure le spin de sa particule suivant $\hat{a}$ et trouve un

spin up, elle sait que si Bob effectue ultérieurement une mesure du spin de sa particule suivant $\hat{b} \equiv \hat{a}$, il va trouver un spin down. Comme les deux mesures sont séparées par un intervalle de genre espace, il est impossible, selon EPR, que le choix de la mesure faite par d'Alice ait pu influencer l'état physique du spin mesuré par Bob, et par conséquent le spin mesuré par Bob devait posséder *avant sa mesure* les informations nécessaires pour réagir comme prévu par la théorie quantique. Autrement dit, le résultat préexiste à la mesure, en contradiction avec les principes de la mesure quantique dont on admet qu'elle ne révèle pas une réalité préexistante. Comme la direction $\hat{a}$ est arbitraire et qu'Alice peut décider au tout dernier moment de changer son axe de mesure, on peut poursuivre l'argumentation : quelle que soit l'orientation $\hat{b}$ choisie par Bob, son spin doit disposer de toute l'information nécessaire pour réagir correctement à sa mesure si l'on veut retrouver les corrélations prévues par la physique quantique. La paire de spins doit donc être créée avec toute l'information nécessaire pour répondre à des mesures suivant des axes arbitraires : les résultats sont prédéterminés par l'état initial de la paire, de la même façon que dans l'exemple du § 11.2.1, l'information sur la couleur des habits enfilés par Alice et Bob est contenue initialement dans les sacs opaques qu'ils ont transportés. Au contraire, ce type d'information n'est pas contenu dans le formalisme de la théorie quantique, et EPR en concluent que la description par la théorie quantique d'une paire de spins *individuelle* est incomplète, même si les prédictions de la théorie quantique sont *statistiquement* correctes.

Continuons notre raisonnement en suivant cette fois Bell [2004], et revenons au diagramme d'espace-temps de la figure 11.6. Considérons une région d'espace-temps P antérieure à R_A et R_B, dont l'intersection avec les cônes passés de R_A et R_B est non-nulle dans une région où ils ne se recouvrent pas. Soit X un ensemble de données (dispositif et procédure de préparation des paires de spins, positionnement des détecteurs...) décrivant complètement l'expérience, à l'exclusion de l'orientation des axes de mesure, et appelons λ l'ensemble des données (nombres, vecteurs...) nécessaires pour compléter la description quantique au sens de EPR : l'ensemble des λ forme les *variables additionnelles*. Les données X et λ sont supposées spécifier *complètement* le processus physique dans l'intersection de la région P avec les cônes passés de R_A et R_B. Soit $\mathsf{p}(A, B|\hat{a}, \hat{b}, X, \lambda)$ la probabilité conditionnelle d'obtenir le résultat joint (A, B) pour une orientation $\hat{a}, \hat{b}$ des axes, probabilité qui dépend aussi *a priori* de X et de λ ; afin d'alléger les notations, nous écrivons simplement a et b au lieu de $\hat{a}$ et $\hat{b}$. Comme la description par (X, λ) est supposée complète, elle est aussi déterministe (voir Nordsen [2006] pour une étude détaillée de ce point) : à un ensemble donné (X, λ) correspond un seul couple de valeurs possibles pour A et B donnant une probabilité non nulle, ce qui implique la factorisation

$$\mathsf{p}(A, B|a, b, X, \lambda) = \mathsf{p}(A|a, b, X, \lambda)\mathsf{p}(B|a, b, X, \lambda)$$

En invoquant maintenant la causalité, nous pouvons simplifier l'expression ci-dessus

$$\mathsf{p}(A, B|a, b, X, \lambda) = \mathsf{p}(A|a, X, \lambda)\mathsf{p}(B|b, X, \lambda) \tag{11.58}$$

En effet, comme R_A et R_B sont séparés par un intervalle de genre espace, $\mathsf{p}(A|a, b, X, \lambda)$ ne peut pas dépendre de l'orientation b choisie par Bob. Introduisons finalement une loi de probabilité $\rho(\lambda, a, b, X)$ pour λ et ajoutons l'hypothèse cruciale suivante : Alice et Bob ont la liberté de choisir au tout dernier moment, c'est-à-dire en particulier hors de la région P, l'orientation de leurs axes de mesure, ce que l'on peut appeler le libre arbitre des expérimentateurs[16]. Si ces choix sont vraiment aléatoires, alors $\rho(\lambda, a, b, X)$ ne peut pas dépendre de a et b. La distribution de probabilité des résultats A et B s'obtient en intégrant sur λ

$$\mathsf{p}(A, B|a, b, X) = \int \mathrm{d}\lambda \rho(\lambda, X) \mathsf{p}(A|a, X, \lambda) \mathsf{p}(B|b, X, \lambda) \qquad (11.59)$$

Cette forme de $\mathsf{p}(A, B|a, b, X)$ permet de retrouver immédiatement l'inégalité de Bell (11.42), car par exemple

$$\langle AB \rangle (X) = \int \mathrm{d}\lambda \rho(\lambda, X)\, AB\, \mathsf{p}(A|a, X, \lambda) \mathsf{p}(B|b, X, \lambda) \qquad (11.60)$$

et on peut utiliser l'égalité (11.41). Si l'on admet que cette inégalité BCHSH est contredite par l'expérience, on en déduit :

(i) qu'il n'est pas possible de "compléter" la mécanique quantique de la façon envisagée par EPR ;

(ii) qu'il n'est pas possible que les corrélations quantiques mesurées dans R_A et R_B aient leur origine uniquement dans les cônes passés de R_A et R_B (figure 11.6).

Les points (i) et (ii) montrent que la physique quantique est incompatible avec la localité, ce qui entraîne l'impossibilité d'une description de certains processus quantiques comme une suite d'événements se déroulant dans l'espace-temps, lorsque le système physique s'étale sur une grande région de l'espace et forme une entité unique[17]. En effet, la mesure "à distance" des spins d'Alice et de Bob est inconciliable avec la notion d'événement par définition ponctuel. Pour pouvoir décrire un tel processus comme résultant d'une suite d'événements dans l'espace-temps, il faudrait que ses résultats découlent de causes communes dans leurs cônes passés, ce qui est exclu par la violation des inégalités de Bell. La physique quantique ne contredit pas la relativité (heureusement !), mais elle en restreint la portée en comparaison du cas classique.

Remarquons en conclusion de cette section que nous avons donné deux démonstrations apparemment indépendantes des inégalités de Bell, l'une fondée sur l'existence d'une distribution de probabilité $\mathsf{p}(A, A', B, B')$ et l'autre

16. Nous excluons le superdéterminisme, l'existence d'une intelligence laplacienne qui pourrait tout connaître du futur, y compris le choix qui sera fait pour les axes. Les expérimentateurs, qui sont en pratique des générateurs de nombres aléatoires, seraient alors entièrement prédéterminés et ne disposeraient pas de leur libre arbitre.

17. Un article de Conway et Kochen [2009] qui utilise des paires de particules massives de spin 1 montre encore plus directement cette propriété.

sur l'expression (11.59). Fine [1982] a montré que les deux hypothèses sont mathématiquement équivalentes : l'existence d'une distribution de probabilité $\mathsf{p}(A, A', B, B')$ implique l'équation (11.59) et réciproquement. Autrement dit, on peut arriver à (11.42) soit en partant de la contrafactualité, soit en partant de la localité. Cependant, la démonstration fondée sur l'argument EPR de localité est la plus intéressante car elle conduit à réviser notre conception de l'espace-temps. Le raisonnement qui mène à (11.60) est souvent justifié par une hypothèse supplémentaire de déterminisme, mais, ainsi que nous l'avons vu, cette hypothèse n'est pas nécessaire : *le déterminisme résulte de la localité*. Enfin, on déduit de l'argumentation précédente que toute théorie de variables additionnelles reproduisant les résultats de la physique quantique doit être non locale. La théorie de ce type la plus développée, celle de de Broglie/Bohm, est effectivement explicitement non locale : elle contient des influences instantanées à distance (voir par exemple Holland [1993]).

11.3 Compléments sur les inégalités de Bell

11.3.1 Conditions sur les probabilités

De façon générale, on peut exprimer les résultats d'une expérience schématisée sur la figure 11.4 en se donnant une distribution de probabilité conditionnelle $\mathsf{p}(A, B|a, b)$, la probabilité d'observer les résultats A et B étant donné les orientations a et b, ce qui donne 16 probabilités différentes. Récrivons tout d'abord l'inégalité BCHSH (11.42) en utilisant des probabilités au lieu de valeurs moyennes, en remarquant que la valeur moyenne $\langle AB\rangle$ peut s'écrire en fonction des probabilités $\mathsf{p}(A = B|a, b)$ et $\mathsf{p}(A \neq B|a, b)$ sous la forme

$$\begin{aligned}\langle AB\rangle &= \mathsf{p}(A = B|a, b) - \mathsf{p}(A \neq B|a, b)\\ &= 2\mathsf{p}(A = B|a, b) - 1 = 1 - 2\mathsf{p}(A \neq B|a, b)\end{aligned} \qquad (11.61)$$

L'inégalité BCHSH s'écrit alors

$$\mathsf{p}(A = B|a, b) + \mathsf{p}(A = B'|a, b') + \mathsf{p}(A' = B|a', b) + \mathsf{p}(A' \neq B'|a', b') \leq 3 \qquad (11.62)$$

En fait, il n'est pas indispensable de préciser les axes qui sont implicites dans l'écriture des résulats, mais nous les avons conservés par souci de cohérence avec ce qui suit. La borne de Cirelson devient $2 + \sqrt{2} \simeq 3.41$ dans le membre de droite de (11.62) et viole bien évidemment l'inégalité BCHSH.

Examinons maintenant les conditions générales auxquelles doivent obéir les probabilités $\mathsf{p}(A, B|a, b)$. La première condition est la normalisation

$$\sum_{A,B} \mathsf{p}(A, B|a, b) = 1 \qquad (11.63)$$

ce qui réduit de 16 à 12 le nombre de probabilités indépendantes. Une condition essentielle est la causalité relativiste : il faut que les résultats d'Alice, par

exemple, ne puissent pas être influencés par l'orientation b choisie par Bob

$$\mathsf{p}(A|a,b) \equiv \sum_B \mathsf{p}(A,B|a,b) = \mathsf{p}(A|a) \tag{11.64}$$

La probabilité $\mathsf{p}(A|a)$ ne dépend que de a, et pas de b. Mais cela n'implique en aucun cas une factorisation

$$\mathsf{p}(A,B|a,b) = \mathsf{p}(A|a)\mathsf{p}(B|b)$$

car une telle factorisation exclurait toute corrélation entre A et B ! Il faut bien comprendre la différence avec la factorisation dans (11.58) : c'est le caractère supposé complet de la description (λ, X) qui entraîne la factorisation à λ fixé, et les corrélations sont introduites par l'intégration sur λ. La condition (11.64) exprime la causalité relativiste et la mécanique quantique obéit à cette causalité. En effet, la probabilité pour Alice de trouver le résultat $A = \pm 1$ dans la mesure de $(\vec{\sigma}\cdot\hat{a})$ est donnée par la valeur moyenne du projecteur $\mathcal{P}_a^A$ sur le sous espace de la valeur propre A

$$\mathcal{P}_a^A = \frac{1}{2}(I + A\,\vec{\sigma}\cdot\hat{a})$$

et $\mathsf{p}(A|a)$ s'écrit lorsque le vecteur d'état des deux spins 1/2 est $|\Psi\rangle$

$$\mathsf{p}(A|a) = \langle\Psi|\mathcal{P}_a^A \otimes I_{\mathcal{B}}|\Psi\rangle = \mathrm{Tr}\,(\rho_{\mathcal{A}}\mathcal{P}_a^A) \tag{11.65}$$

où $I_{\mathcal{B}}$ est l'opérateur identité pour le spin $\mathcal{B}$ et $\rho_{\mathcal{A}}$ l'opérateur statistique réduit du spin $\mathcal{A}$

$$\rho_{\mathcal{A}} = \mathrm{Tr}_{\mathcal{B}}(|\Psi\rangle\langle\Psi|)$$

Le résultat (11.65) ne dépend que de paramètres contrôlés par Alice. Jusqu'à la fin de cette section, il sera commode d'utiliser les notations de l'informatique, auxquelles nous reviendrons dans la section 11.5. Au lieu d'attribuer les valeurs ± 1 à A et B, nous allons choisir les valeurs $A, B = 1$ ou $A, B = 0$, de sorte que A et B prennent les valeurs 0 ou 1 d'un bit classique. De même, au lieu de a et a', nous utiliserons $a = 0$ pour a et $a = 1$ pour a', de sorte que l'orientation des axes est aussi spécifiée par la valeur d'un bit classique. Nous appellerons A_0 et A_1 les résultats correspondant à a_0 et a_1 : $A \to A_0$, $A' \to A_1$. Avec ces notations, il est immédiat de réécrire l'inégalité BCHSH (11.62) sous la forme (Buhrman *et al.* [2010])

$$R = \frac{1}{4}\sum_{a,b} \mathsf{p}(A \oplus B = ab|a,b) \leq \frac{3}{4} \tag{11.66}$$

où $\oplus$ est l'addition modulo 2

$$0 \oplus 0 = 1 \oplus 1 = 0 \qquad 0 \oplus 1 = 1 \oplus 0 = 1 \tag{11.67}$$

On remarque que la condition $A \oplus B = ab$ s'écrit sous forme explicite

$$\begin{aligned} A_0 \oplus B_0 &= 0 \qquad & A_0 \oplus B_1 &= 0 \\ A_1 \oplus B_0 &= 0 \qquad & A_1 \oplus B_1 &= 1 \end{aligned} \tag{11.68}$$

Ces quatre équations sont incompatibles : en prenant leur somme, le membre de gauche vaut zéro, le membre de droite un. Au maximun trois équations sur quatre sont compatibles, ce qui explique intuitivement le facteur 3/4 dans (11.66). Les détails sont renvoyés à l'exercice 11.6.9.

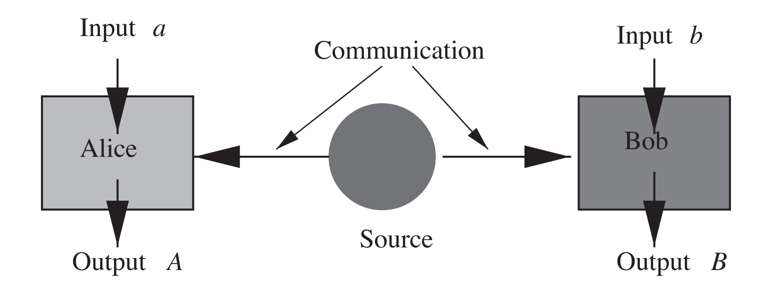


FIG. 11.9 – Schéma de communication et boîtes noires.

Les expériences décrites dans la section 11.2 sont schématisées sur la figure 11.9 : une source S permet à Alice et à Bob de partager une information, qui peut être une information préétablie, un état intriqué ou toute autre forme de connexion imaginable. Alice et Bob possèdent chacun une boîte noire avec une entrée (input) qui reçoit le bit a (Alice) ou b (Bob). La sortie (output) de chaque boîte est le bit A ou B : A_0 si $a = 0$, A_1 si $a = 1$ et de même B_0 si $b = 0$, B_1 si $b = 1$. La stratégie gagnante consiste à obtenir le résultat $A \oplus B = ab$ quels que soient a et b. Avec une stratégie de communication préétablie, le taux de succès de la stratégie optimale est de 75 %, en raison de (11.66), et d'environ 85 % si Alice et Bob partagent un état intriqué : $(2+\sqrt{2})/4 = \cos^2 \pi/8 \simeq 0.853$. Cela résulte de la borne de Cirelson, mais la stratégie optimale peut aussi être déterminée directement (exercice 11.6.9).

11.3.2 Boîtes de Popescu-Rohrlich

Les quatre probabilités dans (11.66) sont bornées par l'unité, et *a priori* le membre de droite R pourrait atteindre la valeur 1, au lieu de 3/4 : communication préétablie ou $(2+\sqrt{2})/4)$: borne de Cirelson. Peut-on atteindre $R = 1$ sans violer la condition essentielle de causalité relativiste ? La réponse est positive, comme l'ont montré Popescu et Rohrlich [1987] qui choisissent

$$\begin{aligned} \mathsf{p}(A \oplus B = ab|a, b) &= \frac{1}{2} \\ \mathsf{p}(A \oplus B \neq ab|a, b) &= 0 \end{aligned} \tag{11.69}$$

Les probabilités non nulles sont toutes égales à 1/2 et s'écrivent explicitement

$$\begin{aligned} \mathsf{p}(0,0|0,0) = \mathsf{p}(1,1|0,0) &= \frac{1}{2} \qquad & \mathsf{p}(0,0|0,1) = \mathsf{p}(1,1|0,1) &= \frac{1}{2} \\ \mathsf{p}(0,0|1,0) = \mathsf{p}(1,1|0,0) &= \frac{1}{2} \qquad & \mathsf{p}(0,1|1,1) = \mathsf{p}(1,0|1,1) &= \frac{1}{2} \end{aligned} \tag{11.70}$$

Reste à vérifier la condition de causalité relativiste, ce qui se fait par inspection, par exemple pour $A = 0 = a = 0$

$$\mathsf{p}(A = 0|0,0) = \mathsf{p}(A = 0|0,1) = \frac{1}{2}$$

Il résulte immédiatement de (11.70) que la somme des probabilités dans (11.62) est égale à 4 et $R = 1$. Si les boîtes noires de la figure 11.9 sont des boîtes de Popescu-Rohrlich vérifiant (11.69), alors la stratégie optimale est gagnante à 100 % ! Les corrélations des boîtes de Popescu-Rohrlich sont encore plus fortes que celles de la mécanique quantique, tout en respectant la condition de causalité relativiste. On peut se demander quelle est la spécficité de la physique quantique qui limite les corrélations à la borne de Cirelson. Plusieurs auteurs ont tenté de répondre à cette question, sans qu'une réponse entièrement satisfaisante n'ait émergé pour l'instant.

11.3.3 États GHZ

Il existe des généralisations intéressantes des états intriqués au cas de trois particules, en particulier les états dits GHZ (Greenberger, Horne et Zeilinger).

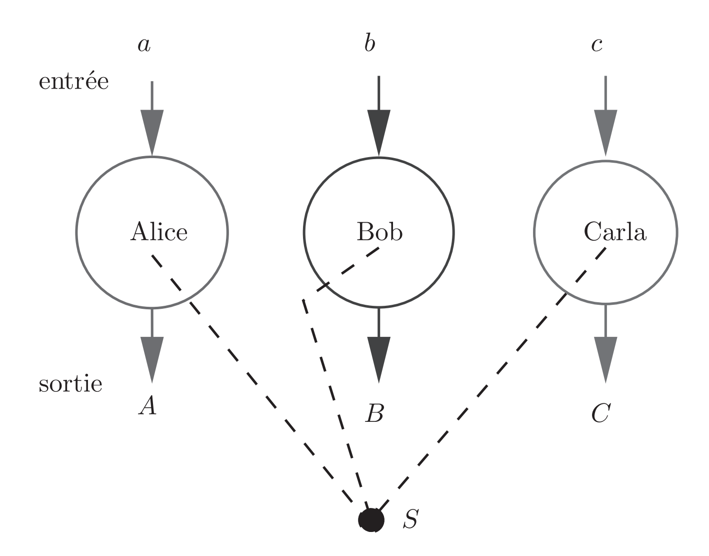


Fig. 11.10 – Cas de trois particules. Alice, Bob et Carla peuvent partager une communication préétablie ou un état intriqué dont la source est S.

Les trois particules sont observées par trois expérimentateurs, Alice ($\mathcal{A}$), Bob ($\mathcal{B}$) et Carla ($\mathcal{C}$). Le schéma de la figure 11.9 est généralisé au cas de trois particules, chaque expérimentateur disposant d'une orientation (ou entrée) a, b ou c et obtient des résultats (ou sorties) A, B ou C ; comme ci-dessus les résultats A_0, B_0 et C_0 correspondent respectivement aux entrées $a = 0$, $b = 0$ et $c = 0$, tandis que les résultats A_1, B_1 et C_1 correspondent aux entrées respectives $a = 1$, $b = 1$ et $c = 1$ (figure 11.10). La stratégie gagnante consiste à obtenir

$$\begin{aligned} &A \oplus B \oplus C = 0 \qquad \text{si} \quad abc = \{000\} \\ &A \oplus B \oplus C = 1 \qquad \text{si} \quad abc \in \{011, 101, 110\} \end{aligned} \tag{11.71}$$

ou de façon explicite

$$\begin{aligned} A_0 \oplus B_0 \oplus C_0 &= 0 \qquad A_0 \oplus B_1 \oplus C_1 = 1 \\ A_1 \oplus B_0 \oplus C_1 &= 1 \qquad A_1 \oplus B_1 \oplus C_0 = 1 \end{aligned}$$

Toute stratégie déterministe où $A_0 \dots C_1$ sont fonctions de $a_0 \dots c_1$ est vouée à l'échec, car les 4 équations sont incompatibles ; la somme des membres de gauche est nulle, alors que celle des membres de droite vaut 1. Il est facile de montrer que toute stratégie de communication préétablie probabiliste est limitéee à un taux de succès de 75 % (exercice 11.6.9). En revanche, il existe une stratégie quantique dont le taux de succès est de 100 % ! Nous allons utiliser non seulement les notations de l'informatique classique, mais aussi celles de l'informatique quantique (section 11.5). Au lieu de noter les états de base $|+\rangle$ et $|-\rangle$, nous les notons $|0\rangle$ et $|1\rangle$, ce qui est la notation usuelle pour des bits quantiques, ou qubits

$$|+\rangle \to |0\rangle \qquad |-\rangle \to |1\rangle \tag{11.72}$$

Supposons qu'Alice, Bob et Carla partagent un état intriqué à trois particules, appelé état GHZ

$$|\Psi\rangle = \frac{1}{2}(|000\rangle - |011\rangle - |101\rangle - |111\rangle) \tag{11.73}$$

Si les trois bits d'entrée sont $abc = \{000\}$, alors les trois expérimentateurs mesurent dans la base $\{|0\rangle, |1\rangle\}$, et on constate sur (11.73) que le résultat est $A_0 \oplus B_0 \oplus C_0 = 0$ dans tous les cas : par exemple, le second terme de (11.73) donne $0 \oplus 1 \oplus 1 = 0$. Si $\{abc\} = \{011\}$, alors Bob et Carla appliquent l'opérateur de Hadamard H, que nous reverrons dans la section 11.5

$$\mathrm{H}|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \qquad \mathrm{H}|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

à leurs états respectifs, avec pour résultat

$$I_{\mathcal{A}} \otimes \mathrm{H}_B \otimes \mathrm{H}_C |\Psi\rangle = \frac{1}{2}(|001\rangle + |010\rangle - |100\rangle + |111\rangle)$$

Quel que soit le terme du second membre, $A_0 \oplus B_1 \oplus C_1 = 1$. L'étude des deux autres cas est renvoyée à l'exercice 11.6.9. Le partage d'un état intriqué fait mieux que toute stratégie de communication préétablie : il permet de passer d'un taux de succès de 75 % à un taux de 100 %. On observe également qu'en passant d'un état intriqué de deux particules à trois particules, on améliore le taux de succès de $\simeq 85$ % à 100 %.

11.3.4 Contextualité

Nous avons vu qu'une théorie de variables additionnelles doit être non locale si l'on exige qu'elle reproduise les réultats de la mécanique quantique. Une autre propriété qu'une telle théorie doit posséder est la *contextualité*. Afin d'expliquer ce concept, supposons que nous ayons en mécanique quantique deux ensembles d'opérateurs hermitiens (ou propriétés physiques) compatibles $\{K, L, M, \ldots\}$ et $\{K, P, Q \ldots\}$, mais que les deux ensembles d'opérateurs ne soient pas compatibles entre eux : $[K, L] = [K, P] = [L, M] \ldots = 0$, tandis que $[L, P] \neq 0$. Pour des opérateurs compatibles, une relation opératorielle $f(K, L, M, \ldots) = 0$ implique que leurs valeurs propres $k, l, m \ldots$ obéissent à la même relation fonctionnelle : $f(k, l, m, \ldots) = 0$. Peut-on généraliser cette propriété si on considère des opérateurs non compatibles ? Pour répondre à cette question, considérons le tableau suivant de 9 opérateurs formés avec les opérateurs de spin de deux particules $\mathcal{A}$ et $\mathcal{B}$ (Mermin [1993])

$$\begin{array}{ccc}
K = \sigma_x^A & L = \sigma_x^B & M = \sigma_x^A \sigma_x^B = KL \\
P = \sigma_y^B & Q = \sigma_y^A & R = \sigma_y^A \sigma_y^B = QP \\
U = KP = \sigma_x^A \sigma_y^B & V = QL = \sigma_y^A \sigma_x^B & W = \sigma_z^A \sigma_z^B = UV
\end{array}$$

Ces 9 opérateurs ont pour valeurs propres ± 1, les 3 opérateurs de chaque ligne commutent et il en est de même pour ceux de chaque colonne. On devrait donc avoir

$$m = kl \qquad r = qp \qquad w = uv$$

On en déduit

$$mrw = (kl)(qp)(uv) = (kp)(lq)(uv) = (uv)^2 = 1$$

Ce résultat est en contradiction avec l'identité opératorielle $MRW = -I$. Nous venons d'exhiber un cas particulier du *théorème de Kochen-Specker* : il est impossible d'associer leurs valeurs propres à chacun des opérateurs d'un sous-ensemble d'opérateurs compatibles, si l'on exige que ces valeurs propres obéissent aux mêmes relations fonctionnelles que les opérateurs. Dans une théorie de variables additionnelles reproduisant les résultats de la mécanique quantique, les valeurs de ces variables sont nécessairement *contextuelles* : elles dépendent en général des propriétés physiques qui sont mesurées simultanément. C'est le cas par exemple de la théorie de de Broglie-Bohm : Holland [1993]. Si on mesure une propriété physique K simultanément à L

et M, on ne peut pas lui attribuer la même valeur que si on la mesure simultanément à P et Q. Hagesawa *et al.* [2006] ont montré qu'une théorie de variables additionnelles non contextuelle était incompatible avec l'expérience. Cependant, l'observation que toute théorie de variables additionnelles doit être contextuelle n'est pas obligatoirement un argument décisif contre de telles théories.

11.4 Décohérence et mesure

11.4.1 Intrication et perte de cohérence

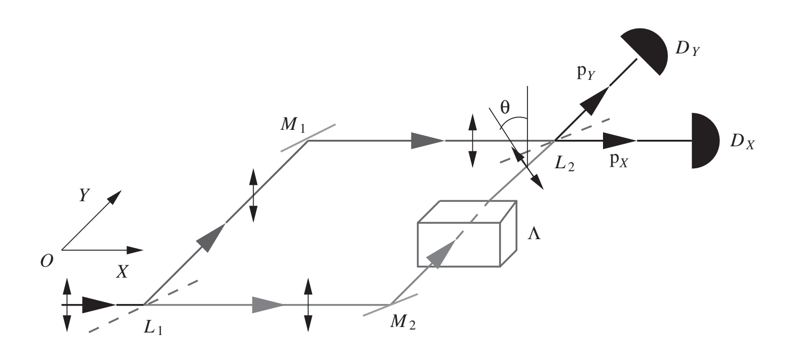


FIG. 11.11 – Un photon unique polarisé verticalement entre dans l'interféromètre en L_1. Entre le miroir M_2 et la lame L_2, une lame biréfringente fait tourner la polarisation de θ, alors que sur le trajet $L_1M_1L_2$ la polarisation reste verticale.

Nous allons introduire la notion de perte de cohérence sur un exemple simple, l'interféromètre de Mach-Zehnder du § 1.4.5 en utilisant cette fois des photons polarisés. À l'entrée de l'interféromètre situé dans le plan XOY (figure 11.11), les photons sont polarisés perpendiculairement à ce plan, et une lame biréfringente Λ fait tourner la polarisation d'un angle θ sur le trajet $L_1M_2L_2$. Les photons se propagent suivant des directions OX ou OY (propriétés spatiales) et de façon à éviter toute confusion avec les degrés de liberté de propagation, nous allons noter V un état de polarisation perpendiculaire au plan de l'interféromètre, H un état de polarisation dans ce plan ; enfin nous supposons les lames L_1 et L_2 équilibrées et sans influence sur la polarisation. Soit a l'amplitude de probabilité à l'entrée de l'interféromètre, que nous prendrons égale à un ultérieurement. À la sortie de la lame L_2, nous devons tenir compte de 4 amplitudes de probabilité, qui se déduisent des résultats du § 1.4.5 pour des lames équilibrées avec $t = 1/\sqrt{2}, r = \mathrm{i}/\sqrt{2}$.

1. Photon sortant de L_2 dans la direction X polarisé verticalement : amplitude de probabilité
   $a_X(V) = \dfrac{\mathrm{i}a}{2}(1 + \cos\theta)$

2. Photon sortant de L_2 dans la direction X polarisé horizontalement : amplitude de probabilité
$a_X(H) = \frac{\mathrm{i}a}{2}\sin\theta$
3. Photon sortant de L_2 dans la direction Y polarisé verticalement : amplitude de probabilité
$a_Y(V) = \frac{a}{2}(-1+\cos\theta)$
4. Photon sortant de L_2 dans la direction Y polarisé horizontalement : amplitude de probabilité
$a_Y(H) = \frac{a}{2}\sin\theta$

Pour calculer les probabilités, nous appliquons les règles du § 1.4.3 en tenant compte de ce que les états de polarisation V et H sont *discernables*. Nous pourrions en effet analyser la polarisation des photons après la lame L_2 à l'aide d'un prisme polarisant et identifier leurs polarisations. Nous avons quatre probabilités différentes notées suivant le même schéma que les amplitudes, à savoir : $\mathsf{p}_X(V), \ldots, \mathsf{p}_Y(H)$. La probabilité p_X pour qu'un photon sorte suivant la direction OX s'obtient en additionnant les contributions des polarisations verticale et horizontale, et on doit additionner les probabilités, pas les amplitudes !

$$\begin{aligned}
\mathsf{p}_X &= \mathsf{p}_X(V) + \mathsf{p}_X(H) = |a_X(V)|^2 + |a_X(H)|^2 = \frac{1}{4}\,(1+\cos\theta)^2 + \frac{1}{4}\,\sin^2\theta \\
&= \frac{1}{2}\,(1+\cos\theta) \\
\mathsf{p}_Y &= \mathsf{p}_Y(V) + \mathsf{p}_Y(H) = |a_Y(V)|^2 + |a_Y(H)|^2 = \frac{1}{4}\,(1-\cos\theta)^2 + \frac{1}{4}\,\sin^2\theta \\
&= \frac{1}{2}\,(1-\cos\theta)
\end{aligned} \tag{11.74}$$

Nous en déduisons la *visibilité des franges*

$$\mathcal{V} = \frac{\mathsf{p}_X - \mathsf{p}_Y}{\mathsf{p}_X + \mathsf{p}_Y} = \cos\theta \tag{11.75}$$

La visibilité des franges est donc réduite par un facteur $\cos\theta$, et elle s'annule si $\theta = \pi/2$: les interférences ont disparu ! L'explication est simple : pour $\theta = \pi/2$, la *polarisation agit comme un marqueur du trajet suivi par le photon* : l'observation d'une polarisation V (H) indique que le photon a suivi le trajet $L_1M_1L_2$ $(L_1M_2L_2)$, et nous pouvons diviser l'ensemble des photons en deux groupes, ceux qui ont suivi le trajet $L_1M_1L_2$ et ceux qui ont suivi le trajet $L_1M_2L_2$ et additionner les contributions des deux trajets. On ne peut pas observer d'interférences car elle sont détruites par l'information sur le trajet.

En combinant les propagations suivant les deux axes OX et OY, la lame L_2 permet de tester la *cohérence* entre ces deux états. La cohérence est maximale si $\mathcal{V} = 1$, c'est-à dire pour $\theta = 0$ et elle est nulle pour $\mathcal{V} = 0$, c'est-à dire

pour $\theta = \pi/2$. En d'autres termes, la lame L_2 permet de tester la superposition linéaire des deux trajets. Cette superposition est cohérente pour $\theta = 0$, incohérente pour $\theta = \pi/2$. Examinons le rôle de l'intrication en utilisant le formalisme de Dirac.

- $|X\rangle$ ($|Y\rangle$) = photon se propageant dans la direction X (Y)
- $|V\rangle$ ($|H\rangle$) = photon polarisé verticalement (horizontalement)

L'état quantique du photon avant la seconde lame est un état intriqué

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(\cos\theta|Y \otimes V\rangle + \sin\theta|Y \otimes H\rangle + |X \otimes V\rangle\right)$$

ce qui donne un opérateur statistique global $\rho_{\text{tot}} = |\Psi\rangle\langle\Psi|$ tandis que l'opérateur statistique spatial du photon s'obtient en prenant la trace partielle sur les degrés de liberté de polarisation

$$\rho_{\text{spat}} = \text{Tr}_{\text{pol}}\,\rho_{\text{tot}} = \frac{1}{2}(I + \cos\theta\,\sigma_x) = \frac{1}{2}\begin{pmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{pmatrix} \tag{11.76}$$

Ajoutons trois remarques importantes.

1. Il n'y a eu aucune perturbation de la propagation spatiale des photons. Il n'est donc pas possible d'expliquer la disparition des interférences par une "perturbation incontrôlable". C'est *l'étiquetage des trajets* par la polarisation qui est à la base de la disparition des interférences.
2. Il n'est pas nécessaire d'observer la polarisation pour que les interférences disparaissent. Il suffit que le dispositif puisse en principe donner accès à cette information. Le comportement du photon n'est pas lié à une observation effectuée ou non effectuée, *mais à la configuration du dispositif expérimental*, une propriété que nous avons déjà observée pour l'expérience à choix retardé.
3. En manipulant le photon après L_2, nous pourrions récupérer les interférences et effacer l'information contenue dans la polarisation. C'est avec un dispositif un peu différent, mais dont le principe est analogue, que fonctionne l'expérience à choix retardé. C'est *uniquement* lorsque l'un des compteurs a été déclenché que nous pouvons affirmer avec certitude que le photon a suivi un des deux trajets. Comme le dit John A. Wheeler : "Aucun phénomène élémentaire ne peut être considéré comme tel tant qu'il n'a pas été enregistré."

En résumé, l'interféromètre de Mach-Zehnder avec photons polarisés illustre un cas où il existe une intrication quantique entre un degré de liberté de propagation (la propagation suivant les directions X ou Y) et un degré de liberté interne, la polarisation. C'est cette intrication d'un degré de liberté de propagation avec un degré de liberté de polarisation qui a permis d'influencer la cohérence du premier. Si l'on observe seulement le degré de liberté de propagation, on constate une perte de cohérence lorsque l'angle θ passe de

0 à $\pi/2$. La perte de cohérence est totale lorsque les états de polarisation intriqués avec chacun des deux états de propagation sont discernables, par exemple les états $|V\rangle$ et $|H\rangle$. Cette perte de la cohérence due à l'intrication est précisément le phénomène de *décohérence*.

11.4.2 Définition générale de la décohérence

Nous allons examiner un système intriqué de $\mathcal{H}_A \otimes \mathcal{H}_E$ formé du système qui nous intéresse, $\mathcal{A}$, et d'un environnement $\mathcal{E}$: l'espace de Hilbert sera donc $\mathcal{H}_A \otimes \mathcal{H}_E$. Dans l'exemple que nous venons de décrire, les degrés de liberté spatiaux décrits par $\{|X\rangle, |Y\rangle\}$ sont ceux qui nous intéressent, tandis que les degrés de liberté internes $\{|V\rangle, |H\rangle\}$ sont ceux de l'environnement. Le phénomème de décohérence est contenu dans l'équation (11.16) dont un cas particulier est (11.76) pour $\theta = \pi/2$: dans l'état intriqué, $|\Phi_{AE}\rangle$ construit avec des vecteurs $|\varphi_{i_A}\rangle \in \mathcal{H}_A$ et les vecteurs $|\chi_{i_E}\rangle$ d'une base orthonormée de $\mathcal{H}_E$

$$|\Phi_{AE}\rangle = \sum_i c_i |\varphi_{i_A} \otimes \chi_{i_E}\rangle \tag{11.77}$$

l'état $|\varphi_{i_A}\rangle \in \mathcal{H}_A$ du système est corrélé à un état $|\chi_{i_E}\rangle \in \mathcal{H}_E$ de l'environnement, et ces états vérifient la condition d'orthogonalité

$$\langle \chi_{i_E} | \chi_{j_E} \rangle = \delta_{ij} \tag{11.78}$$

Dans ces conditions, l'équation (11.34) qui généralise (11.76) pour $\theta = 0$ donne l'opérateur statistique de $\mathcal{A}$

$$\rho_A = \mathrm{Tr}_{\mathcal{E}} |\Phi_{AE}\rangle\langle\Phi_{AE}| = \sum_i |c_i|^2 |\varphi_{i_A}\rangle\langle\varphi_{i_A}| \tag{11.79}$$

ρ_A ne dépend que du module de c_i, et toute l'information sur les phases des nombres complexes c_i a disparu dans ρ_A. En résumé, *si des états du système $\mathcal{A}$ sont corrélés à des états orthogonaux de l'environnement $\mathcal{E}$, toute cohérence de phase entre les états de $\mathcal{A}$ est perdue.* Cette perte de la cohérence de phase est appelée *décohérence*.

Comme dans l'exemple du § 11.4.1, la décohérence vient de ce que l'état de $\mathcal{E}$ agit comme un *marqueur* de l'état de $\mathcal{A}$. Donnons un autre exemple, en imaginant une expérience de fentes d'Young où une particule 1 passe à travers une des fentes, et soit $|a_1\rangle$ (resp. $|a_1'\rangle$) son état quantique lorsqu'elle passe par la fente F (resp. F'), c'est-à-dire son état quantique quand la fente F' (resp. F) est fermée. Supposons l'état de la particule 1 intriqué avec celui d'une particule 2, l'état global $|\Psi\rangle$ étant

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|a_1 \otimes a_2\rangle + |a_1' \otimes a_2'\rangle \right)$$

Si, par exemple, les deux particules sont issues de la désintégration d'une particule instable d'impulsion nulle, leurs impulsions sont corrélées par conserva-

tion de l'impulsion

$$\vec{p}_1 + \vec{p}_2 = 0$$

La mesure de $\vec{p}_2$ donne une information sur $\vec{p}_1$, elle peut permettre sous certaines conditions de remonter à la trajectoire de la particule 1, et par exemple de déterminer la fente d'Young choisie par celle-ci, ce qui entraîne la destruction des interférences.

Comme nous l'avons vu, la décohérence détruit la cohérence de phase. Toutefois, les phases ont seulement disparu *localement*, c'est-à-dire si on se limite à des mesures sur le système $\mathcal{A}$. Il est facile de trouver des propriétés physique jointes de $\mathcal{A}$ et $\mathcal{E}$ dont les éléments de matrice dépendent des phases des coefficients c_i. Si l'on est capable de garder le contrôle de l'ensemble des variables quantiques du système $\mathcal{AE}$, alors on a affaire à une "fausse décohérence", dont un exemple (théorique !) a été donné par Raimond *et al.* [1997]. Un autre exemple est fourni par une expérience de trous d'Young effectuée avec des atomes : si les électrons sont considérés comme l'environnement des noyaux, il est clair que les fonctions d'onde électroniques correspondant aux atomes passant par l'une ou l'autre des deux fentes sont orthogonales. Mais les fonctions d'onde électroniques suivent adiabatiquement le mouvement des noyaux, et la recombinaison des deux amplitudes se fait sur l'écran de façon cohérente, avec des interférences. Au contraire, si l'on perd le contrôle des variables de l'environnement, alors on a affaire à une vraie décohérence : l'information fuit de façon incontrôlable dans l'environnement. Par exemple, si l'environnement est constitué des photons du rayonnement du fond cosmologique à 3 K, il est évidemment impossible de garder le contrôle sur les photons de ce rayonnement qui ont diffusé sur le système.

Le *temps de décohérence* $\tau_{\rm dec}$ est le temps caractéristique de décroissance de la cohérence de phase. Ce temps de décohérence est d'autant plus court que le système $\mathcal{A}$ est complexe. Nous allons le voir sur un exemple simple ; des modèles plus élaborés seront discutés dans le chapitre 18 et confirmeront la validité de cet énoncé. Supposons qu'un spin 1/2 interagisse avec son environnement de telle sorte que rien ne change s'il est dans l'état $|+\rangle$, mais que son vecteur d'état change de signe s'il est dans l'état $|-\rangle$

$$|+\rangle \to |+\rangle \qquad |-\rangle \to -|-\rangle \tag{11.80}$$

Dans un intervalle Δt, le processus (11.80) appelé *basculement de phase*[18] se produit avec une probabilité $\mathsf{p} = \Gamma\Delta t$, où Γ représente la probabilité par unité de temps. Nous supposerons toujours $\mathsf{p} \ll 1$ de telle sorte que la probabilité pour deux basculements de phase dans l'intervalle Δt soit négligeable. Si l'état de spin est la superposition linéaire vecteur propre de σ_x avec la valeur propre $+1$

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$$

18. Ce processus est important dans la discussion des codes correcteurs d'erreur en informatique quantique.

cet état se transforme par basculement de phase en

$$|\psi\rangle \to |\psi'\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)$$

vecteur propre de σ_x avec la valeur propre -1, avec une probabilité $\mathsf{p} = \Gamma\Delta t$: le temps de décohérence est $\tau_{\rm dec} = 1/\Gamma$. Partons maintenant d'un état GHZ de N spins ayant la forme suivante

$$|\Psi_n\rangle = \frac{1}{\sqrt{2}}\big(|++\cdots+\rangle + |--\cdots-\rangle\big) \tag{11.81}$$

Pour que la relation de phase initiale entre les deux composantes de $|\Psi_n\rangle$ soit perdue, il suffit *qu'un seul* des spins soit affecté

$$|\Psi_n\rangle \to |\Psi'_n\rangle = \frac{1}{\sqrt{2}}\big(|++\cdots+\rangle - |--\cdots-\rangle\big) \tag{11.82}$$

En admettant en première approximation que le processus (11.80) se produit indépendamment pour tous les spins, le processus (11.82) a une probabilité $n\mathsf{p} = n\Gamma\Delta t$ de se produire dans l'intervalle de temps Δt. Le temps de décohérence $\tau_{\rm dec}^{(n)} = \tau_{\rm dec}/n$ est d'autant plus court que le nombre de spins est grand. Ceci est un trait général de la décohérence qui sera étudié plus en détail au chapitre 18. Par exemple, dans le mouvement brownien quantique, il suffit d'une seule collision pour provoquer la décohérence, alors que plusieurs collisions sont nécessaires pour changer la vitesse de façon appréciable : pour un système complexe, *le temps de décohérence est extrêmement court par rapport aux autres temps caractéristiques, par exemple le temps de relaxation de l'énergie*. Compte tenu de cette propriété, l'observation expérimentale du phénomène de décohérence n'est pas facile. Cependant, la décohérence a été observée dans des conditions parfaitement contrôlées par Brune *et al.* [1996].

11.4.3 Modèle pour l'émission spontanée

Pour illustrer le phénomène de décohérence, nous allons étudier un modèle très schématique d'émission spontanée d'un photon par un atome. Le système $\mathcal{A}$ est l'atome, qui peut être dans un état excité $|1_A\rangle$ ou dans l'état fondamental $|0_A\rangle$. L'environnement est le champ électromagnétique quantifié, qui peut être dans son état fondamental à zéro photon, $|0_E\rangle$, ou dans un état à un photon $|1_E\rangle$. L'évolution unitaire se fait dans l'espace $\mathcal{H}_A \otimes \mathcal{H}_E$, et si l'état initial est $|0_A \otimes 0_E\rangle$, rien ne peut se passer

$$U_{AE}|0_A \otimes 0_E\rangle = |0_A \otimes 0_E\rangle \tag{11.83}$$

Mais si l'état initial est $|1_A \otimes 0_E\rangle$, l'atome peut émettre dans l'intervalle de temps Δt un photon avec une probabilité $\mathsf{p} = \Gamma\Delta t$, où Γ est la largeur de raie, c'est-à-dire l'inverse de la vie moyenne de l'état excité

$$U_{AE}|1_A \otimes 0_E\rangle = \sqrt{1-\mathsf{p}}\;|1_A \otimes 0_E\rangle + \sqrt{\mathsf{p}}\;|0_A \otimes 1_E\rangle \tag{11.84}$$

Il est instructif de choisir comme état initial de l'atome une superposition linéaire des états $|0_A\rangle$ et $|1_A\rangle$, tandis que l'état du champ est à zéro photon, $|0_E\rangle$

$$|\Phi_{AE}\rangle = (\lambda|0_A\rangle + \mu|1_A\rangle) \otimes |0_E\rangle \tag{11.85}$$

ce qui correspond à la matrice statistique suivante pour $\mathcal{A}$

$$\rho_A = \begin{pmatrix} |\lambda|^2 & \lambda\mu^* \\ \lambda^*\mu & |\mu|^2 \end{pmatrix}$$

Utilisons (11.83)-(11.84) pour calculer la matrice statistique $\rho_A^{(1)}$ au temps Δt. Le vecteur d'état devient l'état intriqué

$$U_{AE}|\Phi_{AE}\rangle = \left(\lambda|0_A\rangle + \mu\sqrt{1-\mathsf{p}}\;|1_A\rangle\right) \otimes |0_E\rangle + \mu\sqrt{\mathsf{p}}\;|0_A \otimes 1_E\rangle \tag{11.86}$$

et la matrice $\rho_A^{(1)}$

$$\rho_A^{(1)} = \mathrm{Tr}_{\mathcal{E}}\left[U_{AE}|\Phi_{AE}\rangle\langle\Phi_{AE}|U_{AE}^\dagger\right] = \begin{pmatrix} 1-(1-\mathsf{p})|\mu|^2 & \sqrt{1-\mathsf{p}}\;\lambda\mu^* \\ \sqrt{1-\mathsf{p}}\;\lambda^*\mu & (1-\mathsf{p})|\mu|^2 \end{pmatrix}$$

Pour obtenir la matrice statistique au temps t, on divise t en n intervalles $\Delta t = t/n$, avec $\Gamma\Delta t \ll 1$

$$\rho_A^{(n)} = \begin{pmatrix} 1-(1-\mathsf{p})^n\rho_{11} & (1-\mathsf{p})^{n/2}\,\rho_{01} \\ (1-\mathsf{p})^{n/2}\,\rho_{10} & (1-\mathsf{p})^n\rho_{11} \end{pmatrix} \to \begin{pmatrix} 1-\mathrm{e}^{-\Gamma t}\,\rho_{11} & \mathrm{e}^{-\Gamma t/2}\,\rho_{01} \\ \mathrm{e}^{-\Gamma t/2}\,\rho_{10} & \mathrm{e}^{-\Gamma t}\,\rho_{11} \end{pmatrix} \tag{11.87}$$

On observe que le temps de décroissance des cohérences est $T_2 = 2/\Gamma$, et celui des populations est $T_1 = 1/\Gamma$, et donc $T_2 = 2T_1$. Ce résultat, qui nous servira au chapitre 15, est spécifique du cas où l'environnement est limité au seul champ électromagnétique. Dans le cas d'environnements plus complexes (avec des collisions, etc.), les cohérences dans la base {état fondamental, état excité} sont en général beaucoup plus fragiles que les populations, et $T_2 \ll T_1$. C'est aussi la situation générale en RMN dans la base $\{|+\rangle, |-\rangle\}$ pour un champ $\vec{B}_0$ orienté suivant Oz, et c'est pourquoi on utilise une impulsion $\pi/2$ de préférence à une impulsion π : le retour à l'équilibre se fait plus rapidement.

Une seconde observation est la suivante : supposons que le détecteur de photons ait une efficacité de 100 %, qu'il couvre tout l'angle solide, et que nous n'observions aucun photon entre les instants $t = 0$ et $t = \Delta t$. Cela veut dire qu'en partant de l'état (11.85) $\lambda|0_A\rangle + \mu|1_A\rangle$ de l'atome, nous l'avons préparé au temps Δt dans l'état (non normalisé)

$$\lambda|0_A\rangle + \mu\sqrt{1-\mathsf{p}}\;|1_A\rangle \tag{11.88}$$

Bien que nous n'ayons absolument pas agi sur l'atome, son vecteur d'état a évolué en raison de l'absence de détection d'un photon ! Il ne peut donc y avoir de “perturbation incontrôlable due à la mesure”, et on observe aussi le rôle de la non localité dans la mesure où l'atome et le détecteur sont localisés en des régions d'espace différentes.

11.4.4 Modèle de von Neumann pour la mesure

Une application très importante de l'intrication est la théorie de la mesure quantique due à von Neumann. Dans cette théorie, on suppose que la mesure doit avoir comme phase initiale une interaction quantique entre le système étudié et l'appareil de mesure, considéré, au moins au départ, comme un objet quantique. L'appareil de mesure $\mathcal{M}$ est donc traité comme un système quantique, et on se propose d'effectuer des mesures sur un système $\mathcal{A}$. Soit $|\varphi_n\rangle$, une base de $\mathcal{H}_A$ formée de vecteurs propres d'une propriété physique A, les valeurs propres a_n étant supposées non dégénérées pour simplifier la discussion

$$A|\varphi_n\rangle = a_n|\varphi_n\rangle \tag{11.89}$$

Soit $|\Phi_n\rangle$ une base orthonormée de $\mathcal{H}_M$, $\langle\Phi_n|\Phi_m\rangle = \delta_{nm}$, et soit $|\Phi_0\rangle$ l'état initial de l'appareil de mesure, qui est *a priori* quelconque, pas nécessairement un des états $|\Phi_n\rangle$. L'état initial, avant toute interaction entre $\mathcal{A}$ et $\mathcal{M}$, est un produit tensoriel $|\Psi_i\rangle$

$$|\Psi_i\rangle = |\varphi_n \otimes \Phi_0\rangle \tag{11.90}$$

et le résultat de l'interaction est de transformer $|\Psi_i\rangle$ en $|\Psi_f\rangle$

$$|\Psi_i\rangle \to |\Psi_f\rangle = U(\infty)|\Psi_i\rangle = |\varphi_n \otimes \Phi_n\rangle \tag{11.91}$$

L'évolution (11.91) est bien compatible avec l'unitarité : elle transforme des états orthogonaux en états orthogonaux, $\langle\Psi_i|\Psi'_i\rangle = 0 \Longrightarrow \langle\Psi_f|\Psi'_f\rangle = 0$. Un modèle explicite qui réalise cette évolution est décrit dans l'exercice 8.6.13. L'équation (11.91) montre que si $\mathcal{M}$ est observé dans l'état $|\Phi_n\rangle$, cela implique que $\mathcal{A}$ est dans l'état $|\varphi_n\rangle$, et le résultat de la mesure de A est a_n. Si l'état initial de $\mathcal{A}$ est une superposition linéaire $\sum_n c_n|\varphi_n\rangle$, soit

$$|\Psi_i\rangle = \left(\sum_n c_n|\varphi_n\rangle\right) \otimes |\Phi_0\rangle \tag{11.92}$$

alors $|\Psi_f\rangle$ est un état intriqué

$$|\Psi_f\rangle = \sum_n c_n|\varphi_n \otimes \Phi_n\rangle \tag{11.93}$$

Dans ces conditions, l'opérateur statistique de $\mathcal{A}$ est alors

$$\rho_A = \mathrm{Tr}_{\mathcal{M}}|\Psi_f\rangle\langle\Psi_f| = \sum_n |c_n|^2|\varphi_n\rangle\langle\varphi_n| \tag{11.94}$$

Un examen superficiel de (11.94) pourrait conduire à la conclusion suivante : nous venons de montrer la règle de Born et la réduction du paquet d'ondes. En effet, $\mathcal{A}$ est trouvé dans l'état $|\varphi_n\rangle$ avec la probabilité $|c_n|^2$. Bien entendu, c'est une illusion, car en premier lieu la prescription de trace partielle utilisée pour

démontrer (11.94) a été déduite de la règle de Born, et la "démonstration" est en fait un raisonnement circulaire. En second lieu, il n'y a pas eu réduction du paquet d'ondes, car le mélange dans (11.94) est un mélange impropre (§ 11.1.4), et non un mélange incohérent d'états $|\varphi_n\rangle$ avec la probabilité $|c_n|^2$. Mais il y a plus grave. Un premier problème est celui de l'ambiguïté de la propriété physique mesurée, que nous allons illustrer en prenant pour $\mathcal{A}$ et $\mathcal{M}$ deux spins 1/2, une base $|z_\pm\rangle$ d'états propres de σ_z pour $\mathcal{A}$ et une base $|Z_\pm\rangle$ d'états propres de Σ_z pour $\mathcal{M}$: afin d'alléger les notations, nous notons *uniquement jusqu'à la fin de cette section* les vecteurs de base $|z_\pm\rangle$ au lieu de $|\pm, z\rangle$. Nous aurons également besoin des bases de vecteurs propres $|x_\pm\rangle = (|z_+\rangle \pm |z_-\rangle)/\sqrt{2}$ de σ_x dans $\mathcal{H}_A$ et $|X_\pm\rangle = (|Z_+\rangle \pm |Z_-\rangle)/\sqrt{2}$ de Σ_x dans $\mathcal{H}_M$. Nous choisissons $|\Phi_0\rangle = |X_+\rangle$ et l'évolution suivante

$$\begin{aligned} U(\infty)|z_+ \otimes X_+\rangle &= |z_+ \otimes Z_+\rangle \\ U(\infty)|z_- \otimes X_+\rangle &= |z_- \otimes Z_-\rangle \end{aligned} \tag{11.95}$$

Partant d'une superposition linéaire initiale

$$|\Psi_i\rangle = (\lambda|z_+\rangle + \mu|z_-\rangle) \otimes |X_+\rangle$$

nous arrivons à l'état final

$$|\Psi_f\rangle = \lambda|z_+ \otimes Z_+\rangle + \mu|z_- \otimes Z_-\rangle \tag{11.96}$$

Comme on s'y attendait, les phases de λ et μ restent présentes dans les éléments de matrice de propriétés jointes de $\mathcal{A}$ et $\mathcal{M}$, par exemple

$$\langle\Psi_f|\sigma_x\Sigma_x|\Psi_f\rangle = \frac{1}{2}\Big[|\lambda|^2 + |\mu|^2 + 2\mathrm{Re}(\lambda^*\mu)\Big] \tag{11.97}$$

et de plus, si nous choisissons $\lambda = \mu = 1/\sqrt{2}$, (11.96) peut se récrire

$$|\Psi_f\rangle = \frac{1}{\sqrt{2}}\Big(|x_+ \otimes X_+\rangle + |x_- \otimes X_-\rangle\Big) \tag{11.98}$$

ce qui fait que $\mathcal{M}$ mesure σ_x, et non σ_z : il y a ambiguïté sur la propriété physique mesurée. En observant $\mathcal{M}$ longtemps après que la mesure a été effectuée, l'expérimentateur peut choisir de mesurer σ_x ou σ_z, qui sont pourtant des propriétés physiques incompatibles.

Le deuxième problème est celui du chat de Schrödinger. Revenant au cas général, si l'état initial de $\mathcal{A}$ est une combinaison linéaire

$$|\varphi\rangle = \lambda|\varphi_1\rangle + \mu|\varphi_2\rangle$$

alors l'état final est

$$|\Psi_f\rangle = \lambda|\varphi_1 \otimes \Phi_1\rangle + \mu|\varphi_2 \otimes \Phi_2\rangle \tag{11.99}$$

Si $\mathcal{M}$ est macroscopique, on obtient une superposition linéaire de deux états macroscopiquement distincts, appelée *chat de Schrödinger* : cette dénomination vient d'une situation imaginée par Schrödinger en 1935 afin d'en souligner l'absurdité, et est parfois appelée paradoxe du chat de Schrödinger. Un chat est enfermé dans une boîte contenant aussi un noyau atomique radioactif, qui se désintègre en moyenne au bout d'une heure. Si le noyau se désintègre, un dispositif diabolique tue le malheureux chat : une particule émise dans la désintégration déclenche un compteur, lui-même connecté à marteau qui brise une fiole contenant un gaz instantanément mortel. L'état du chat est intriqué avec celui de l'atome comme dans (11.99), et au bout d'une heure, l'état quantique est une superposition à poids égaux de l'atome non désintégré $|\mathrm{ATOME}\rangle$ et du chat vivant $|\mathrm{CHAT}\rangle$, et de l'atome désintégré $|\mathrm{atome}\rangle$ et du chat mort $|\mathrm{chat}\rangle$

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\mathrm{ATOME} \otimes \mathrm{CHAT}\rangle + |\mathrm{atome} \otimes \mathrm{chat}\rangle)$$

La théorie de von Neumann a pour conséquence inévitable que si l'on doit traiter les appareils macroscopiques comme des objets quantiques, on est alors conduit à des états de superposition qui semblent *a priori* absurdes et qui n'ont jamais été observés. En résumé, la théorie de von Neumann doit faire face à deux problèmes : (i) l'équation (11.99) ne décrit pas une mesure effectivement réalisée, mais seulement ce qui est souvent appelé une pré-mesure et (ii) la théorie introduit inévitablement des superpositions d'états macroscopiquement discernables. Les deux composantes de $|\Psi_f\rangle$ sont présentes dans l'état final, on ne peut pas dire que $\mathcal{M}$ est dans l'état $|\Phi_1\rangle$ ou dans l'état $|\Phi_2\rangle$ et aucune mesure n'a été effectuée, car aucune des éventualités : $\mathcal{M}$ dans l'état $|\Phi_1\rangle$ ou $\mathcal{M}$ dans l'état $|\Phi_2\rangle$ n'a été réalisée : *la théorie de von Neumann ne permet pas de faire émerger un résultat unique.*

On pourrait compliquer le schéma de von Neumann en ajoutant un appareil de mesure $\mathcal{M}_1$ qui mesure $\mathcal{M}$, puis $\mathcal{M}_2$ qui mesure $\mathcal{M}_1$, etc. : l'ensemble est appelé *chaîne de von Neumann*. Mais tant que l'évolution est linéaire et régie par l'équation de Schrödinger, on bute sur le même problème : la mesure ne fait pas émerger un résultat unique. Il faut donc, à un certain moment, briser la chaîne de von Neumann à l'étape i, en admettant qu'un des appareils $\mathcal{M}_i$ se retrouve dans un état unique et bien défini, autrement dit un état classique. C'est ce qu'admettent Landau et Lifschitz [1966], qui commencent par montrer l'équation (11.93) et ajoutent : “le caractère classique de l'appareil s'exprime dans le fait que, à chaque instant, on peut affirmer qu'il se trouve dans un des états $|\Phi_n\rangle$”, un énoncé parfaitement arbitraire et parachuté d'autorité. von Neumann brise la chaîne au niveau de l'observateur. C'est la lecture de l'appareil par l'observateur qui inscrit de façon indélébile le résultat de la mesure dans son cerveau, et c'est alors que se passe le processus d'enregistrement irréversible de la mesure. Cet aspect de la théorie de von Neumann n'est plus pris au sérieux aujourd'hui, ne serait-ce que parce que l'enregistrement des clics des détecteurs se fait de façon automatique avec pilotage par

un ordinateur et les cerveaux humains n'interviennent que bien après que les résultats ont été enregistrés de façon irréversible. L'observation n'intervient que dans la lecture de l'écran ou des listings de données. De même, le décès (éventuel) du chat de Schrödinger ne se produit pas au moment où l'observateur ouvre la boîte, mais au moment où le détecteur enregistre de façon irréversible la particule émise dans la désintégration de l'atome et déclenche le dispositif diabolique. Malgré tout, la première partie de la théorie de von Neumann, celle qui conduit à l'équation (11.93), reste aujourd'hui la base des approches modernes de la mesure.

11.4.5 Modèle de Zurek

Zeh [1970] et Zurek [1991] ont remarqué que $\mathcal{M}$ étant macroscopique, il interagit fortement avec son environnement $\mathcal{E}$. En raison de cette interaction, il existe dans l'espace de Hilbert des états privilégiés, les *états pointeurs*[19], qui ne s'intriquent pas avec l'environnement et qui sont déterminés par la forme du hamiltonien d'interaction H_{ME}. En revanche, toute superposition linéaire de tels états est très rapidement détruite par la décohérence et l'appareil de mesure ne peut "se trouver" que dans un des états pointeurs. Des modèles explicites de ce phénomène sont étudiés au chapitre 18. Cette observation est très importante pour le processus de mesure, car elle permet de lever l'ambiguïté sur la base mentionnée précédemment : si les états $|Z_\pm\rangle$ de (11.95) sont des états pointeurs, les superpositions linéaires

$$|X_\pm\rangle = \frac{1}{\sqrt{2}}\left(|Z_+\rangle \pm |Z_-\rangle\right)$$

ne le sont pas. Cependant, ainsi que nous allons le montrer, la prise en compte de l'environnement ne résout pas le problème de la mesure, ce que suggèrent les guillemets entre lesquels nous avons mis "se trouver".

Dans le modèle de Zurek, $\mathcal{A}$ et $\mathcal{M}$ sont des spins 1/2 comme dans la section 11.4.4, et l'environnement $\mathcal{E}$ est constitué de $N \gg 1$ spins 1/2, $k = 1, 2, \ldots, N$. Le hamiltonien H_{ME} est supposé de la forme

$$H_{ME} = \sum_{k=1}^{N} H_{ME}^{(k)} \qquad H_{ME}^{(k)} = \hbar g_k \Sigma_z \otimes \sigma_z^{(k)} \tag{11.100}$$

où g_k est une constante de couplage sans dimension prenant des valeurs aléatoires en fonction de k ; $|z_\pm^{(k)}\rangle$ désigne les vecteurs propres de $\sigma_z^{(k)}$

$$\sigma_z^{(k)}|z_\pm^{(k)}\rangle = \pm|z_\pm^{(k)}\rangle$$

Après interaction entre $\mathcal{A}$ et $\mathcal{M}$, le vecteur d'état est donné par (11.96)

$$|\Psi_{\rm tot}(0)\rangle = |\Psi_f\rangle \otimes \prod_{k=1}^{N}\Big(\alpha_k|z_+^{(k)}\rangle + \beta_k|z_-^{(k)}\rangle\Big) \tag{11.101}$$

19. En anglais : "pointer states".

Afin de simplifier l'argument, nous avons supposé que le temps caractéristique de l'interaction $\mathcal{AM}$ est très court par rapport à celui de l'interaction $\mathcal{ME}$. Il est facile de déterminer l'évolution ultérieure de $|\Psi_{\rm tot}\rangle$ en remarquant que celle des quatre vecteurs de base suivants est évidente

$$\begin{array}{rcl} |Z_+ \otimes z_+^{(k)}\rangle & \to & {\rm e}^{-{\rm i}g_k t}\,|Z_+ \otimes z_+^{(k)}\rangle \\ |Z_+ \otimes z_-^{(k)}\rangle & \to & {\rm e}^{{\rm i}g_k t}\,|Z_+ \otimes z_-^{(k)}\rangle \\ |Z_- \otimes z_+^{(k)}\rangle & \to & {\rm e}^{{\rm i}g_k t}\,|Z_- \otimes z_+^{(k)}\rangle \\ |Z_- \otimes z_-^{(k)}\rangle & \to & {\rm e}^{-{\rm i}g_k t}\,|Z_- \otimes z_-^{(k)}\rangle \end{array}$$

On obtient donc

$$\begin{array}{rcl} U(t,0)|\Psi_{\rm tot}(0)\rangle & = & \lambda|z_+ \otimes Z_+\rangle \otimes \displaystyle\prod_{k=1}^{N}\left(\alpha_k {\rm e}^{-{\rm i}g_k t}\,|z_+^{(k)}\rangle + \beta_k {\rm e}^{{\rm i}g_k t}\,|z_-^{(k)}\rangle\right) \\ & + & \mu|z_- \otimes Z_-\rangle \otimes \displaystyle\prod_{k=1}^{N}\left(\alpha_k {\rm e}^{{\rm i}g_k t}\,|z_+^{(k)}\rangle + \beta_k {\rm e}^{-{\rm i}g_k t}\,|z_-^{(k)}\rangle\right) \\ & = & \lambda|z_+ \otimes Z_+ \otimes \mathcal{E}_+(t)\rangle + \mu|z_- \otimes Z_- \otimes \mathcal{E}_-(t)\rangle \end{array}$$

Les états $|\mathcal{E}_+\rangle$ et $|\mathcal{E}_-\rangle$ sont corrélés respectivement à $|z_+ \otimes Z_+\rangle$ et à $|z_- \otimes Z_-\rangle$. On en déduit l'opérateur statistique de $\mathcal{AM}$ en prenant la trace partielle sur l'environnement

$$\begin{array}{rcl} \rho_{AM} & = & {\rm Tr}_{\mathcal{E}}\rho_{\rm tot} = {\rm Tr}_{\mathcal{E}}\left[U(t,0)|\Psi_{\rm tot}(0)\rangle\langle\Psi_{\rm tot}(0)|U^\dagger(t,0)\right] \\ & = & |\lambda|^2|z_+ \otimes Z_+\rangle\langle z_+ \otimes Z_+| + |\mu|^2|z_- \otimes Z_-\rangle\langle z_- \otimes Z_-| \\ & + & \left[z(t)\lambda\mu^*|z_+ \otimes Z_+\rangle\langle z_- \otimes Z_-| + {\rm h.c.}\right] \end{array}$$

où $z(t)$ mesure le recouvrement de $|\mathcal{E}_+\rangle$ et de $|\mathcal{E}_-\rangle$

$$z(t) = \langle\mathcal{E}_-(t)|\mathcal{E}_+(t)\rangle = \prod_{k=1}^{N}\left[\cos 2g_k t + {\rm i}(|\alpha_k|^2 - |\beta_k|^2)\sin 2g_k t\right] \qquad (11.102)$$

La superposition d'un grand nombre de fréquences différentes dans (11.102) entraîne que $|z(t)| \to 0$, ce que Zurek a vérifié numériquement, de sorte que

$$\rho_{AM} \to |\lambda|^2|z_+ \otimes Z_+\rangle\langle z_+ \otimes Z_+| + |\mu|^2|z_- \otimes Z_-\rangle\langle z_- \otimes Z_-| \qquad (11.103)$$

Les états $|\mathcal{E}_+\rangle$ et $|\mathcal{E}_-\rangle$ deviennent orthogonaux à la limite $t \to \infty$ et les états $|Z_\pm\rangle$ sont bien des états pointeurs. Alors que le système $\mathcal{AM}$ était décrit dans la section 11.4.4 par un vecteur d'état, la prise en compte de l'environnement fait que ce système est maintenant décrit par un opérateur statistique, dont les cohérences ont été éliminées. Un examen superficiel de la situation pourrait conduire à la conclusion suivante : il n'y a plus de superposition linéaire,

et donc plus d'ambiguïté de la base de mesure, et il n'y a plus de chats de Schrödinger. Le système $\mathcal{AM}$ "se trouve" dans l'état $|z_+ \otimes Z_+\rangle$ avec une probabilité $|\lambda|^2$ et dans l'état $|z_- \otimes Z_-\rangle$ avec une probabilité $|\mu|^2$.

Cependant, est-ce vraiment là la solution au problème de la mesure ? Le consensus qui se dégage aujourd'hui est le suivant : si on se limite à un point de vue strictement opérationnel, le modèle de l'environnement est satisfaisant, mais en tant que théorie fondamentale, il n'est pas pertinent[20]. Tout d'abord, on observe que le problème des superpositions linéaires a simplement été déplacé du système $\mathcal{AM}$ vers le système global $\mathcal{AME}$. La valeur moyenne de la propriété physique suivante (d'Espagnat [1995])

$$R = \sigma_x \otimes \Sigma_x \otimes \prod_{k=1}^{N} \sigma_x^{(k)}$$

est indépendante du temps

$$\langle R \rangle = 2\mathrm{Re}(\lambda^* \mu) \prod_{k=1}^{N} 2\mathrm{Re}(\alpha_k^* \beta_k)$$

et par conséquent les phases de λ, μ, α_k et β_k restent pertinentes. On peut rétorquer que R est en pratique certainement très difficile, voire impossible à mesurer (on est en présence de décohérence vraie), et c'est la raison pour laquelle le modèle de Zurek peut être considéré comme satisfaisant d'un point de vue strictement opérationnel : les phases sont là, mais on ne peut pas les atteindre. Il est cependant peu satisfaisant d'invoquer la difficulté *pratique* d'une mesure comme argument en faveur d'une théorie que l'on voudrait *fondamentale*. La seconde objection au modèle de Zurek (non indépendante de la précédente) est que le mélange (11.103) est un mélange impropre : on ne peut pas dire que le système $\mathcal{AM}$ se trouve dans l'état $|z_+ \otimes Z_+\rangle$ avec une probabilité $|\lambda|^2$ et dans l'état $|z_- \otimes Z_-\rangle$ avec une probabilité $|\mu|^2$, et que cet état nous est simplement inconnu. Si tel était le cas, la mesure ne ferait que révéler un résultat qui nous était précédemment caché, car nous aurions affaire à un mélange propre (*cf.* la discussion du § 11.1.4). Mais (11.103) décrit un mélange impropre, aucune des deux éventualités $|z_+ \otimes Z_+\rangle$ ou $|z_- \otimes Z_-\rangle$ n'est en fait réalisée ... à moins de disposer d'un appareil $\mathcal{M}'$ extérieur à $\mathcal{AM}$ et qui permet d'observer ce système. Mais si nous supposons l'existence d'un tel appareil, nous sommes ramenés au problème précédent pour le système $\mathcal{AMM}'$! En conclusion, la prise en compte de la décohérence ne permet pas d'expliquer l'émergence d'un résultat unique.

Dans le cas d'une expérience d'interférences du type fentes d'Young, réalisée par exemple avec des atomes, on ne peut pas dire que l'éventualité : l'atome numéro n est passé par la fente F_1 ou l'atome numéro n est passé par la fente

20. "Point de vue strictement opérationnel" est la traduction de "For all practical purposes" (FAPP) (Bell [1990]), et "non pertinent" celle de "It is a non starter" (Leggett [2002b]).

F_2, a été réalisée. Si tel était le cas, on n'observerait pas d'interférences. Si les superpositions d'états macroscopiquement distincts sont possibles, alors il n'y a pas plus de raison pour que l'éventualité : l'ensemble (système + appareil de mesure) est dans l'état $|z_+ \otimes Z_+\rangle$, ou l'ensemble (système + appareil de mesure) est dans l'état $|z_- \otimes Z_-\rangle$, soit réalisée. La théorie quantique est une "théorie totalitaire" (Leggett [2008]), qui englobe en principe aussi bien le microscopique que le macroscopique, et on ne peut pas en changer subrepticement les règles en passant du microscopique au macroscopique !

Est-ce à dire que la décohérence est un concept sans intérêt ? Tel n'est pas le cas, car la décohérence est fondamentale pour expliquer pourquoi un objet macroscopique nous *apparaît* classique, parce que nous sommes en pratique limités à des observations locales. L'observation d'un objet macroscopique nous donne une image classique, exempte de superpositions linéraires, mais la décohérence ne nous permet pas de dire que l'objet "se trouve" dans tel ou tel état.

11.4.6 La réduction du paquet d'ondes

Revenons maintenant sur une question que nous avions laissée en suspens au chapitre 4, celle de la "réduction du paquet d'ondes"[21]. Imaginons que nous effectuions au temps t_1 sur un système $\mathcal{A}$ une mesure d'une propriété physique A (11.89) et au temps $t_2 > t_1$ une mesure d'une propriété physique B dont les valeurs propres sont b_q et les vecteurs propres $|\psi_q\rangle$

$$B|\psi_q\rangle = b_q|\psi_q\rangle$$

Afin de nous débarrasser de l'évolution temporelle des vecteurs d'état, nous nous plaçons dans le point de vue de Heisenberg, et les propriétés physiques mesurées sont en fait $A(t_1)$ et $B(t_2)$. Simplifions la discussion sans affecter la physique sous-jacente en supposant que les mesures s'effectuent pendant un temps très court $\sim \Delta t$ par rapport au temps caractéristique d'évolution de $A(t)$ et $B(t)$, dans deux intervalles de temps $[t_1-\Delta t, t_1+\Delta t]$ et $[t_2-\Delta t, t_2+\Delta t]$ qui ne se recouvrent pas. Le rôle de $U(\infty)$ dans (11.91) est joué par deux opérateurs unitaires U_1 et U_2

$$\begin{aligned} U_1|\varphi_n \otimes \Phi_0\rangle &= |\varphi_n \otimes \Phi_n\rangle \\ U_2|\psi_q \otimes \Psi_0\rangle &= |\psi_q \otimes \Phi_q\rangle \end{aligned}$$

où $|\Phi\rangle$ et $|\Psi\rangle$ sont des vecteurs de l'espace des états $\mathcal{H}_{M_1}$ et $\mathcal{H}_{M_2}$ des appareils de mesure $\mathcal{M}_1$ et $\mathcal{M}_2$. Si l'état initial de $\mathcal{A}$ est $|\varphi\rangle = \sum_n c_n|\varphi_n\rangle$, l'action de

21. Rappelons que contrairement à une opinion largement répandue, la réduction du paquet d'ondes ne fait pas partie de l'interprétation de Copenhague, du moins dans la version de Bohr. Cette réduction du paquet d'ondes n'a évidemment aucun sens dans l'interprétation minimaliste (§ 11.4.7), qui raisonne uniquement sur des ensembles. Les (excellents) livres de Leslie Ballentine [1998] et de Asher Peres [1993] sont deux exceptions notables qui échappent à l'orthodoxie des manuels.

U_1 sur l'état initial de $\mathcal{H}_A \otimes \mathcal{H}_{M_1} \otimes \mathcal{H}_{M_2}$ est

$$U_1|\varphi \otimes \Phi_0 \otimes \Psi_0\rangle = \sum_n \langle\varphi_n|\varphi\rangle|\varphi_n \otimes \Phi_n \otimes \Psi_0\rangle$$

car U_1 (U_2) agit dans $\mathcal{H}_{M_1}$ ($\mathcal{H}_{M_2}$). L'action de U_2 se traduit par

$$U_2U_1|\varphi \otimes \Phi_0 \otimes \Psi_0\rangle = \sum_{n,q} \langle\varphi_n|\varphi\rangle\langle\psi_q|\varphi_n\rangle|\varphi_n \otimes \Phi_n \otimes \Psi_q\rangle$$

D'après le postulat **II** (règle de Born), cette équation donne immédiatement la probabilité jointe

$$\mathsf{p}(B = b_q; A = a_n) = |\langle\varphi_n|\varphi\rangle|^2\,|\langle\psi_q|\varphi_n\rangle|^2 \tag{11.104}$$

La probabilité conditionnelle d'observer b_q sachant que a_n a été observé dans la première mesure est donnée par la loi de Bayes

$$\mathsf{p}(B = b_q|A = a_n) = \frac{\mathsf{p}(B = b_q; A = a_n)}{\mathsf{p}(A = a_n)} = |\langle\psi_q|\varphi_n\rangle|^2 \tag{11.105}$$

Si nous avions appliqué le "postulat" de réduction du paquet d'ondes, nous aurions conclu qu'après la première mesure, le système $\mathcal{A}$ se trouvait dans l'état (4.7)

$$|\varphi_n\rangle = \frac{\mathcal{P}_n|\varphi\rangle}{\langle\varphi_n|\mathcal{P}_n|\varphi\rangle}$$

et que la probabilité de mesurer b_q était $|\langle\psi_q|\varphi_n\rangle|^2$, en accord avec (11.105). Le "postulat" RPO est donc une *conséquence de la règle de Born*. L'argument ci-dessus montre en outre que la validité du postulat RPO est soumise au caractère idéal de la mesure. Il ne serait pas valable si la loi d'évolution du système $\mathcal{A}$ lors de la première mesure était

$$U_1|\varphi_n \otimes \Phi_0\rangle = |\varphi'_n \otimes \Phi_n\rangle$$

c'est-à-dire un état intriqué. En résumé, le postulat RPO est certes commode, mais il n'est pas fondamental : ce qui l'est vraiment est la règle de Born.

11.4.7 Interprétations

L'absence d'une théorie entièrement satisfaisante de la mesure quantique a pour conséquence que la théorie quantique est souvent "interprétée", sans que ces interprétations aient une influence quelconque sur son utilisation pratique. La première interprétation, la plus dépouillée, pourrait être qualifiée de "minimaliste". Selon cette interprétation, une procédure de préparation permet d'obtenir un ensemble d'objets quantiques tous dans le même état. Il s'ensuit une évolution quantique régie par l'équation d'évolution (4.11), et après un

certain temps, on effectue une mesure sur l'état obtenu après évolution. Il s'agit donc de répéter les mêmes mesures sur une série d'objets quantiques tous préparés dans des conditions identiques, ce qui est schématisé sur la figure 4.2. La théorie permet en principe de calculer la probabilité $\mathsf{p}(D_i)$ que le détecteur D_i soit déclenché, et si l'on a préparé une série de N états identiques, le nombre de clics du détecteur D_i sera en moyenne $N \times \mathsf{p}(D_i)$. Cette interprétation des probabilités de la théorie quantique est manifestement de type fréquentiste.

Dans cette interprétation minimaliste[22], les concepts de la théorie quantique (amplitudes de probabilité, vecteurs d'état...) ne sont que des outils de calcul et ne représentent aucune "réalité" externe ; *le vecteur d'état ne représente rien d'autre que la procédure de préparation*. La théorie ne s'applique qu'à des ensembles et pas à des systèmes quantiques individuels. Ce point de vue, parfois attribué à Einstein qui, nous l'avons vu, considérait de toute façon la théorie quantique comme incomplète, était assez naturel lorsque l'on ne savait manipuler expérimentalement que des ensembles d'objets quantiques : le caractère probabiliste pouvait raisonnablement être attribué à la nécessité d'expérimenter sur des ensembles, par exemple un ensemble de plusieurs milliards d'atomes d'une vapeur de sodium enfermée dans un tube de verre. Il devient difficile à soutenir aujourd'hui : comment ne pas prendre explicitement en considération l'état individuel d'un ion piégé unique que l'on peut observer pendant des heures effectuant des transitions entre différents niveaux d'énergie ? Cependant cet argument, s'il montre que la position minimaliste n'est pas très naturelle dans cette situation, ne permet pas de la rejeter définitivement. En effet, on peut interpréter une série d'observations sur un ion unique comme une suite de mesures effectuées sur des états quantiques identiques.

L'interprétation de Copenhague[23], élaborée entre 1925 et 1927 principalement par Bohr et Heisenberg, applique la théorie quantique à des systèmes individuels, et non à des ensembles[24]. En particulier, *les probabilités de la théorie quantique s'appliquent à des systèmes individuels* et on doit adopter une interpétation bayesienne de ces probabilités. Initialement, cette interpré-

---

22. Interprétation minimaliste et chat de Schrödinger. Pour un tenant de l'interprétation minimaliste, il n'y a strictement aucun problème (sauf peut-être avec la Société Protectrice des Animaux !) L'expérience est faite sur un grand nombre (ensemble) de chats, et au bout d'une heure la moitié des chats sont vivants et la moitié sont morts. La théorie quantique calcule correctement la probabilité de chaque éventualité, et c'est tout ce que l'on peut lui demander. Cela n'a aucun sens de se poser la question du sort du chat numéro 36. Le paradoxe apparaît uniquement quand on applique la théorie quantique à des systèmes individuels.

23. Ou plus exactement, comme l'écrit justement Leggett (Leggett [2002b]), la "non interprétation", car Bohr se refuse à attribuer une quelconque "réalité" aux vecteurs d'état, opérateurs, etc. : "Microscopic entities are not even to be thought as possessing properties in the absence of specification of a macroscopic arrangement".

24. Il n'existe pas de version canonique de cette interpétation. La position de Bohr a évolué au cours du temps, et elle différait notablement sur certains points de celle de Heisenberg. En exagérant un peu, on peut dire que chaque physicien intéressé par le sujet définit sa propre version de l'interprétation de Copenhague. Entrer dans le détail nécessiterait donc plusieurs centaines de pages. Voir par exemple Howard [2004].

tation repose sur deux piliers : les inégalités de Heisenberg et la *complémentarité*. Dans sa version la plus élémentaire, la complémentarité affirme qu'une expérience peut, soit mettre en évidence le caractère particulaire (ou corpusculaire) d'un objet quantique, soit son caractère ondulatoire, mais pas les deux à la fois : les aspects ondulatoire et particulaire sont dits complémentaires. Bien que la complémentarité soit toujours présente dans nombre de manuels récents, il semble que, sous cette forme, le concept n'ait plus guère d'intérêt aujourd'hui et que la "dualité onde-corpusule" soit devenue un concept obsolète (Kaiser et al. [2012]). Enfin, selon Bohr et Heisenberg, le caractère probabiliste vient de ce qu'au moment de la mesure, l'objet quantique interagit avec un appareil macroscopique qui le perturbe de façon aléatoire et incontrôlable : c'est la célèbre "perturbation incontrôlable au moment de la mesure" que l'on trouve encore largement citée aujourd'hui. Les résultats des mesures doivent être exprimés en termes classiques, les seuls qui puissent être transmis de façon intelligible. Bohr et Heisenberg diffèrent sur le rôle de l'observateur auquel Heisenberg, contrairement à Bohr, attribue un rôle central.

En 1935, l'article EPR conduisit Bohr à réviser sa position en profondeur. En effet, nous avons vu que, grâce à la non localité, on peut parfaitement mesurer une propriété physique d'un objet quantique sans aucunement interagir avec celui-ci : dans le processus de mesure, la perturbation ne joue donc pas un rôle fondamental. Cette absence de perturbation est évidente dans les commentaires suivant (11.88), et également dans l'expérience d'interférométrie décrite au § 11.4.1, où le trajet du photon est déterminé grâce à sa polarisation, sans aucunement perturber la trajectoire dans l'espace ; d'autres exemples de mesure sans perturbation, ou mesures non destructrices, ont été proposés récemment et même réalisés expérimentalement pour certains d'entre eux.

La position de Bohr est exposée en particulier dans sa réponse à l'article EPR ([Bohr 1935]). En premier lieu, Bohr réaffirme que l'appareillage servant à expérimenter sur des phénomènes quantiques doit être *décrit* en termes classiques, car seul un vocabulaire utilisant des termes définis en physique classique peut nous permettre de communiquer à d'éventuels interlocuteurs ce que nous avons fait. Par exemple, nous pouvons leur communiquer les positions des détecteurs qui ont enregistré des photons avec leur temps d'arrivée. Cela n'entraîne pas nécessairement que l'appareillage lui-même fonctionne suivant les principes de la physique classique : nous reviendrons ultérieurement sur ce point délicat. Nous avons déjà rencontré le deuxième point de Bohr dans la discussion du § 11.4.1 : un phénomène quantique ne peut être considéré comme tel que lorsque l'expérience est terminée, que ses résultats ont été enregistrés grâce à des processus irréversibles, sans possibilité de retour en arrière, et peuvent être communiqués, nous l'avons vu, en termes classiques.

Il en résulte, et c'est le troisième point de la réponse de Bohr à EPR, que les propriétés d'un système quantique sont indissociables de la configuration expérimentale utilisée pour l'observer. Ce troisième point est au cœur

de l'argumentation de Bohr : les conditions expérimentales constituent un élément inhérent de la description de tout phénomène auquel on peut attacher le terme "réalité physique". *La notion de "réalité physique" d'un système quantique ne peut être dissociée de l'appareillage macroscopique utilisé pour l'observer,* et on ne peut pas faire comme si cette réalité avait une existence indépendante de celui-ci. Pour Bohr, il n'existe pas de propriétés intrinsèques d'un système quantique, et notre représentation de son état quantique par un vecteur d'état ne peut donc être qu'une représentation symbolique, qui ne correspond à aucune réalité externe.

Enfin, et c'est le quatrième point, le concept de complémentarité repose sur la notion de configurations expérimentales mutuellement exclusives. Dans l'expérience de la figure 11.3, on peut choisir, par exemple, les orientations $\hat{a} \parallel \hat{b} \parallel Oz$ ou $\hat{a} \parallel \hat{b} \parallel Ox$. Ces deux configurations expérimentales donnent un exemple de configurations complémentaires, qui sont mutuellement exclusives. Les bases de polarisation $\{|+,z\rangle, |-,z\rangle\}$ ou $\{|+,x\rangle, |-,x\rangle\}$ sont complémentaires. Dans toute situation concrète, des configurations expérimentales mutuellement exclusives, ou complémentaires, permettent une description complète et non ambiguë de l'ensemble des phénomènes quantiques en des termes classiques. Ainsi définie, la complémentarité est le fondement du caractère exhaustif de la description quantique, et la critique par EPR de son caractère incomplet n'est pas fondée.

Le point de vue de Bohr impose de tracer une ligne de démarcation entre le système quantique à étudier et le dispositif expérimental. Cette nécessité de distinguer de façon nette l'appareil de mesure du système étudié constitue selon Bohr la différence principale entre physique quantique et physique classique, car dans cette dernière cette distinction n'existe pas : la description d'un système classique est "intrinsèque", elle ne dépend pas de l'environnement expérimental. Il n'est pas *a priori* nécessaire que le système quantique soit microscopique et l'appareil macroscopique, il ne s'agit donc pas obligatoirement d'une frontière classique/quantique et il se pourrait que l'appareil de mesure ait une partie microscopique. Il existe selon Bohr une "zone tampon" où les descriptions classique et quantique se recouvrent. Cependant, la nécessité logique de faire une distinction de principe entre objet quantique étudié et appareil de mesure classique ne peut pas être formalisée : la question de la limite classique du monde quantique est loin d'être comprise et formalisée de façon satisfaisante. Les recherches théoriques et expérimentales de ces trente dernières années montrent au contraire que la notion de zone tampon invoquée par Bohr est une notion qui reste floue et la frontière classique/quantique est un concept mal défini ne découlant pas, au moins aujourd'hui, de la théorie. Si la mesure commence par une interaction quantique, où doit-on mettre la frontière qui distingue l'objet à étudier de l'appareil de mesure ? Et en fin de compte, on ne peut pas échapper au constat que l'appareil de mesure est fondamentalement régi par des lois quantiques et que sa description classique ne peut être qu'une approximation. On voit mal en pratique comment on pour-

rait éviter d'imposer par décret, comme le font Landau et Lifschitz [1966], une ligne de démarcation entre un monde classique et un monde quantique, et la nécessité de faire appel à cette frontière reste un aspect peu convaincant de la vision de Bohr. Cela suggère que la théorie quantique n'est peut-être pas encore sous sa forme finale.

11.5 Information quantique

Pour conclure ce chapitre, nous allons examiner quelques applications des états intriqués à *l'information quantique*, c'est-à-dire la théorie du traitement et de la transmission de l'information utilisant les spécificités de la mécanique quantique : principe de superposition et intrication. Un ingrédient important de l'information quantique est le *théorème de non-clonage quantique* que nous allons montrer comme résultat préliminaire. Nous examinerons ensuite le calcul quantique, la téléportation et l'échange d'intrication.

11.5.1 Théorème de non-clonage quantique

La condition indispensable pour que la méthode de cryptographie quantique décrite au § 3.1.3 soit parfaitement sûre est que l'espionne Ève ne puisse pas reproduire (cloner) l'état de la particule envoyée par Bob à Alice tout en conservant pour elle le résultat de sa mesure, ce qui rendrait l'interception du message indétectable. Que ceci ne soit pas possible est garanti par le théorème de non-clonage quantique. Pour montrer ce théorème, supposons que l'on souhaite dupliquer un état quantique *inconnu* $|\chi_1\rangle \in \mathcal{H}_A$. Le système sur lequel on veut imprimer la copie est noté $|\varphi\rangle \in \mathcal{H}_B$, où $d_A = d_B$: $|\varphi\rangle$ est l'équivalent de la feuille blanche. Par exemple, si l'on veut cloner un état de spin 1/2 $|\chi_1\rangle$, $|\varphi\rangle$ est aussi un état de spin 1/2. Afin de donner le traitement le plus général possible, nous supposons que nous disposons d'un système quantique auxiliaire[25] dans un état $|m\rangle \in \mathcal{H}_C$. L'évolution quantique la plus générale dans $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ est donnée par un opérateur unitaire U, et nous souhaitons que cet opérateur ait l'action hypothétique suivante

$$U|\chi_1 \otimes \varphi \otimes m\rangle = |\chi_1 \otimes \chi_1 \otimes m(\chi_1)\rangle \tag{11.106}$$

de façon à créer un clone de $|\chi_1\rangle$. Pour un autre original $|\chi_2\rangle$, on devrait avoir

$$U|\chi_2 \otimes \varphi \otimes m\rangle = |\chi_2 \otimes \chi_2 \otimes m(\chi_2)\rangle \tag{11.107}$$

Pour montrer qu'un tel opérateur U ne peut pas exister, évaluons le produit scalaire

$$X = \langle \chi_2 \otimes \varphi \otimes m|U^\dagger U|\chi_1 \otimes \varphi \otimes m\rangle \tag{11.108}$$

de deux façons différentes.

25. En anglais : “ancilla”.

1. Utilisant $U^\dagger U = I$

$$X_1 = \langle \chi_2 \otimes \varphi \otimes m | \chi_1 \otimes \varphi \otimes m \rangle = \langle \chi_2 | \chi_1 \rangle$$

2. Utilisant (11.106) et (11.107)

$$X_2 = \langle \chi_2 \otimes \chi_2 \otimes m(\chi_2) | \chi_1 \otimes \chi_1 \otimes m(\chi_1) \rangle = (\langle \chi_2 | \chi_1 \rangle)^2 \langle m(\chi_2) | m(\chi_1) \rangle$$

Si $\langle \chi_2 | \chi_1 \rangle \neq 0$, la condition $X_1 = X_2$ implique $\chi_1 \equiv \chi_2$ en raison de l'inégalité de Schwartz. Si une machine à cloner peut cloner deux états orthogonaux, elle ne peut pas cloner leurs superpositions linéaires. Cette preuve du théorème de non-clonage explique pourquoi on ne peut pas se restreindre en cryptographie quantique à une base d'états de polarisation orthogonaux $\{|x\rangle, |y\rangle\}$ pour les photons. C'est l'utilisation de superpositions linéaires des états de polarisation $|x\rangle$ et $|y\rangle$ qui permet de détecter la présence éventuelle d'un espion.

Si le clonage quantique était possible, on pourrait communiquer à des vitesses supérieures à celle de la lumière (communication supraluminique). Pour le montrer sur un exemple simple, supposons qu'Alice et Bob partagent des paires de spins 1/2 intriqués dans l'état (11.43), et que Bob mesure σ_x ou σ_z : s'il mesure σ_z, Alice recevra le spin $\mathcal{A}$ dans un des deux états $|\pm\rangle$, et s'il mesure σ_x dans l'un des deux états $|\pm, x\rangle = (|+\rangle \pm |-\rangle)/\sqrt{2}$. Supposons qu'Alice soit capable de faire une copie de l'état qu'elle reçoit : avec sa machine à cloner, qui, rappelons-le, fonctionne quel que soit l'état (inconnu) qu'elle reçoit, elle fabrique l'état $| + \otimes + \rangle$ si elle reçoit $|+\rangle$ et l'état $|+, x \otimes +, x\rangle$ si elle reçoit $|+, x\rangle$. Si Bob mesure σ_z, elle construit donc l'opérateur statistique

$$\rho_z = \frac{1}{2}\big(| + \otimes + \rangle\langle + \otimes + | + | - \otimes - \rangle\langle - \otimes - |\big)$$

et si Bob mesure σ_x, elle construit

$$\rho_x = \frac{1}{2}\big(|+, x \otimes +, x\rangle\langle +, x \otimes +, x| + |-, x \otimes -, x\rangle\langle -, x \otimes -, x|\big)$$

On vérifie aisément (par exemple en calculant $\langle + - | \rho_{x,z} | + - \rangle$) que $\rho_x \neq \rho_z$. Si le clonage était possible, Alice pourrait connaître instantanément le système d'axes utilisé par Bob pour sa mesure, même si elle en est distante de plusieurs années-lumière.

Le clonage parfait étant impossible, on peut se demander s'il est possible de cloner de façon approchée, et, dans ce cas, trouver la meilleure approximation possible d'un état $|\chi\rangle$ compatible avec les règles de la mécanique quantique. Il nous faut donc un critère permettant de définir la “proximité” de deux états quantiques. Il n'existe pas de critère unique, mais le plus utilisé est la *fidélité* $\mathcal{F}$: si l'on essaie d'approcher un état quantique $|\chi\rangle$ par un opérateur statistique ρ, alors la fidélité est définie par

$$\mathcal{F} = \langle \chi | \rho | \chi \rangle = \mathrm{Tr}\,(|\chi\rangle\langle\chi|\rho) \tag{11.109}$$

On note que $\mathcal{F} = 1$ si $\rho = |\chi\rangle\langle\chi|$, et $\mathcal{F} = 0$ si $\rho = |\varphi\rangle\langle\varphi|$ avec $\langle\varphi|\chi\rangle = 0$, c'est-à-dire si $|\varphi\rangle$ est orthogonal à $|\chi\rangle$.

Le cas le plus étudié, et le plus important en pratique, est celui du spin 1/2. Il sera commode, dans la suite de cette section, d'utiliser les notations de l'information quantique déjà introduites en (11.72)

$$|+\rangle \to |0\rangle \qquad |-\rangle \to |1\rangle \tag{11.110}$$

Le clonage optimal a été trouvé par Buzek et Hillery [1996] (BH). On peut toujours choisir comme page blanche $|\varphi\rangle = |0\rangle$, et le clonage approché BH fonctionne en prenant aussi un spin 1/2 comme système $\mathcal{C}$ (les $\otimes$ ont été supprimés pour alléger l'écriture)

$$U|\chi 00\rangle = \sqrt{\frac{2}{3}}\,|\chi\chi\chi_\perp\rangle - \sqrt{\frac{1}{6}}\,\big(|\chi\chi_\perp\rangle + |\chi_\perp\chi\rangle\big)|\chi\rangle \tag{11.111}$$

où $|\chi_\perp\rangle$ est l'état orthogonal à $|\chi\rangle$

$$|\chi\rangle = \lambda|0\rangle + \mu|1\rangle \qquad |\chi_\perp\rangle = -\mu^*|0\rangle + \lambda^*|1\rangle \tag{11.112}$$

L'opérateur statistique de $\mathcal{A}$ s'obtient par une trace partielle

$$\rho_A = \mathrm{Tr}_{\mathcal{BC}}\big[U|\chi 00\rangle\langle\chi 00|U^\dagger\big]$$

Si $|\chi\rangle$ est caractérisé par son vecteur de Bloch $\vec{b}$, l'opérateur statistique de $\mathcal{A}$ sera (exercice 6.6.10)

$$\rho_A = \frac{5}{6}\,|\chi\rangle\langle\chi| + \frac{1}{6}\,|\chi_\perp\rangle\langle\chi_\perp| = \frac{1}{2}\left(I + \frac{2}{3}\,\vec{b}\right) = \rho_B \tag{11.113}$$

La première des équations (11.113) montre que la fidélité du clonage est $\mathcal{F} = 5/6$. La preuve que cette fidélité est optimale, est assez longue, et le lecteur est renvoyé à l'article original de Buzek et Hillery [1996]. On peut montrer qu'une fidélité meilleure que 5/6 permettrait la communication supraluminique : on peut donc voir la borne supérieure $\mathcal{F} \le 5/6$ comme une condition de compatibilité entre la mécanique quantique et la relativité.

Le clonage approché BH est optimal et universel : $\mathcal{F}$ ne dépend pas de l'état $|\chi\rangle$ que l'on souhaite cloner. Il est également symétrique : $\rho_A = \rho_B$. Il est possible d'obtenir une meilleure fidélité si on se limite à une classe particulière de vecteurs de $\mathcal{H}_A$. Un cas important pour la sécurité de la cryptographie quantique est le clonage d'états de la forme

$$|\chi(\phi)\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + \mathrm{e}^{\,\mathrm{i}\phi}|1\rangle\big) \tag{11.114}$$

Les deux bases complémentaires utilisées en cryptographie quantique correspondent à des vecteurs de base construits avec $(\phi = 0, \phi = \pi)$ et $(\phi = \pi/2, \phi = 3\pi/2)$, c'est-à-dire aux vecteurs propres de σ_x et σ_y

$$|\pm x\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle \pm |1\rangle\big) \qquad |\pm y\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle \pm \mathrm{i}|1\rangle\big) \tag{11.115}$$

Ces quatre états appartiennent à un équateur de la sphère de Poincaré-Bloch, et on se pose la question du meilleur clonage possible pour les états de cet équateur, dont l'opérateur statistique est

$$\rho(\phi) = |\chi(\phi)\rangle\langle\chi(\phi)| = \frac{1}{2}\big(I + \cos\phi\,\sigma_x + \sin\phi\,\sigma_y\big) \tag{11.116}$$

On peut montrer que le meilleur clonage possible se passe de spin auxiliaire et fonctionne de la façon suivante

$$\begin{aligned} |0_A 0_B\rangle &\to |0_A 0_B\rangle \\ |1_A 0_B\rangle &\to \cos\eta\,|1_A 0_B\rangle + \sin\eta\,|0_A 1_B\rangle \end{aligned} \tag{11.117}$$

avec $0 \le \eta \le \pi/2$. C'est bien une transformation unitaire dans $\mathcal{H}_A \otimes \mathcal{H}_B$, qui transforme des états orthogonaux en états orthogonaux. On calcule aisément les fidélités (exercice 11.6.11) avec pour résultat

$$\mathcal{F}_A = \frac{1}{2}(1 + \cos\eta) \qquad \mathcal{F}_B = \frac{1}{2}(1 + \sin\eta) \tag{11.118}$$

Si ce clonage est utilisé par Ève, son meilleur choix possible est $\mathcal{F}_A = \mathcal{F}_B$: elle renvoie à Bob un état aussi proche que possible de l'original, tout en gardant pour elle une copie aussi proche que possible de l'original, ce qui correspond à un clonage symétrique plus performant que le clonage universel BH

$$\mathcal{F}_{\rm sym} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) \simeq 0.8535 > \frac{5}{6} \tag{11.119}$$

11.5.2 Calcul quantique

Le second sujet de cette section est le calcul quantique. En théorie de l'information, l'unité élémentaire est le bit, qui peut prendre deux valeurs, conventionnellement 0 ou 1. Ce bit est codé classiquement par un système à deux états, par exemple un condensateur qui peut être non chargé (valeur 0 du bit) ou chargé (valeur 1 du bit). Un bit d'information implique typiquement le transfert de 10^4 à 10^5 électrons dans la mémoire vive d'un ordinateur actuel. Une question intéressante est alors : est-il possible de coder l'information sur des électrons (ou d'autres particules) isolé(e)s ? Ainsi que nous l'avons déjà vu, un système quantique à deux états est susceptible de coder un bit d'information : par exemple, nous avons utilisé au § 3.1.3 deux états de polarisation orthogonaux d'un photon pour coder un bit. Pour fixer les idées, nous allons plutôt utiliser les deux états de polarisation d'un spin 1/2 : par convention l'état de spin up $|+\rangle$ correspondra à la valeur 0, l'état de spin down $|-\rangle$ à la valeur 1, et, suivant (11.110), $|+\rangle \equiv |0\rangle$, $|-\rangle \equiv |1\rangle$. Mais, contrairement au système classique qui ne peut exister que dans les états 0 ou 1, le système quantique peut exister dans des états de superposition linéaire $|\varphi\rangle$ de $|0\rangle$ et $|1\rangle$

$$|\varphi\rangle = \lambda|0\rangle + \mu|1\rangle \qquad |\lambda|^2 + |\mu|^2 = 1 \tag{11.120}$$

Au lieu d'un bit ordinaire, le système quantique code un *bit quantique*, ou *qubit*, dont la valeur dans l'état (11.120) reste indéterminée jusqu'à la mesure de la composante z du spin : la mesure donnera le résultat 0 avec une probabilité $|\lambda|^2$, et le résultat 1 avec une probabilité $|\mu|^2$, ce qui n'est pas en soi une propriété particulièrement utile. L'information codée à l'aide de qubits est un exemple d' information quantique. Le théorème de non-clonage implique qu'il est impossible de recopier cette information exactement.

Supposons que nous voulions inscrire dans un registre un nombre entre 0 et 7. Il faudra pour cela disposer de 3 bits. En effet, dans un système de base 2, on peut représenter un nombre de 0 à 7 par une suite de trois nombres 0 ou 1. Un registre classique codera *une* des 8 configurations suivantes

$$\begin{array}{rclccc} 0 &=& \{000\} & 1=\{001\} & 2=\{010\} & 3=\{011\} \\ 4 &=& \{100\} & 5=\{101\} & 6=\{110\} & 7=\{111\} \end{array}$$

Un système de trois spins 1/2 $\mathcal{A}$, $\mathcal{B}$ et $\mathcal{C}$ permettra également de coder un nombre de 0 à 7, par exemple en faisant correspondre ces nombres aux 8 états suivants de trois spins

$$\begin{array}{cccc} 0:|000\rangle & 1:|001\rangle & 2:|010\rangle & 3:|011\rangle \\ 4:|100\rangle & 5:|101\rangle & 6:|110\rangle & 7:|111\rangle \end{array} \tag{11.121}$$

La notation $|101\rangle$ par exemple est un abrégé pour $|1_A \otimes 0_B \otimes 1_C\rangle$. On notera $|x\rangle$, $x = 0, \ \ldots, \ 7$ un des huit états de (11.121), par exemple $|5\rangle = |101\rangle$, et la base de $\mathcal{H}^{\otimes 3}$ formée des vecteurs $|x\rangle$ est appelée *base de calcul*. Comme on peut former une superposition linéaire des états (11.121), on pourrait en conclure que le vecteur d'état d'un sytème de trois spins nous a permis de coder d'un seul coup $2^3 = 8$ nombres, et avec n spins on pourrait coder 2^n nombres ! Cependant, une mesure des trois spins suivant l'axe Oz donnera nécessairement un des huit états (11.121). Nous disposons d'une importante information virtuelle, mais lorsque nous cherchons à la matérialiser dans une mesure effective, nous ne faisons pas mieux que le système classique : la mesure donne un des huit nombres, et pas les huit à la fois.

Les opérations effectuées par un ordinateur quantique sont des transformations unitaires (4.14) dans l'espace de Hilbert des états $\mathcal{H}^{\otimes n}$ des qubits. Ces opérations sont exécutées au moyen de *portes logiques quantiques*. Il est possible de montrer que toutes les opérations unitaires dans $\mathcal{H}^{\otimes n}$ peuvent être décomposées en

- des transformations sur des qubits *individuels* ;
- des portes dites *control-not* (cNOT) agissant sur des *paires* de qubits, définies en (11.125).

Une porte logique individuelle couramment utilisée est la *porte de Hadamard* (figure 11.12a)

$$\mathrm{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{11.122}$$

dont l'action sur les vecteurs de base $|0\rangle$ et $|1\rangle$ est

$$\mathrm{H}|0\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big) \qquad \mathrm{H}|1\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big) \tag{11.123}$$

En dehors de facteurs de phase, la porte de Hadamard représente l'action d'une lame séparatrice. En appliquant une porte H sur chacun des n qubits dans l'état $|0\rangle$, nous obtenons la combinaison linéaire suivante $|\Phi\rangle$ d'états de la base de calcul

$$|\Phi\rangle = \mathrm{H}^{\otimes n}|0\ldots 0\rangle = \mathrm{H}^{\otimes n}|0^{\otimes n}\rangle = \frac{1}{2^{n/2}}\sum_{x=0}^{2^n-1}|x\rangle \tag{11.124}$$

La porte cNOT (figure 11.12b) a l'action suivante sur un état à deux qubits : si le premier qubit, appelé *qubit de contrôle*, est dans l'état $|0\rangle$, le second qubit, appelé *qubit cible*, est inchangé. Si le qubit de contrôle est dans l'état $|1\rangle$, alors les deux états de base du qubit cible sont échangés : $|0\rangle \leftrightarrow |1\rangle$. La représentation matricielle de la porte cNOT, dans la base $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, est

$$\mathrm{cNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & \sigma_x \end{pmatrix}. \tag{11.125}$$

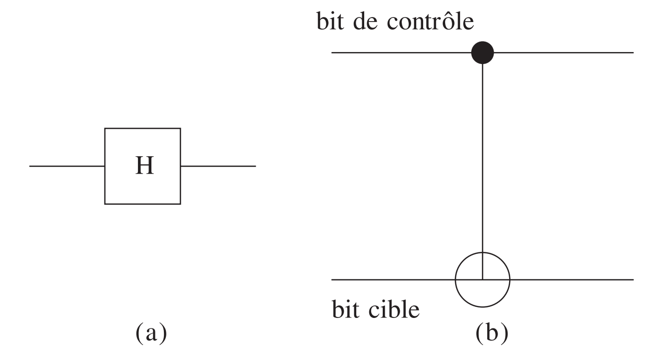


FIG. 11.12 – Représentation graphique des portes logiques quantiques. (a) Porte de Hadamard. (b) Porte cNOT.

Quel gain peut-on attendre d'un ordinateur quantique qui fonctionnerait avec des qubits ? En fait, un ordinateur quantique serait susceptible de mener en parallèle un grand nombre d'opérations. Les opérations élémentaires sur les qubits sont des évolutions unitaires régies par l'équation d'évolution (4.11), ou sa version intégrale (4.14). *Dans certains cas*, une information utile peut être extraite de ces opérations, si l'on utilise le calcul parallèle quantique. Le

principe d'un tel calcul est le suivant : un *registre de données* à n qubits stocke un état $|\Phi\rangle$ (11.124). On construit ensuite le produit tensoriel $|\Psi\rangle$ de $|\Phi\rangle$ avec l'état $|0^{\otimes m}\rangle$ d'un *registre de résultats* à 2^m qubits

$$|\Psi\rangle = |\Phi \otimes 0^{\otimes m}\rangle = \frac{1}{2^{n/2}} \sum_x |x \otimes 0^{\otimes m}\rangle \tag{11.126}$$

Étant donné une fonction $f(x)$ prenant 2^m valeurs différentes stockées dans le registre de résultats, on construit une opération unitaire[26] U_f telle que

$$U_f|x \otimes y\rangle = |x \otimes [y \oplus f(x)]\rangle$$

où $\oplus$ est l'addition modulo 2 sans retenue. Il est clair que $U_f^2 = I$, et que U_f, qui est une simple permutation des vecteurs de base est une opération unitaire. Si U_f est appliqué sur $|\Psi\rangle$ (11.126), le résultat est

$$|\Psi\rangle \to |\Psi'\rangle = U_f|\Psi\rangle = \frac{1}{2^{n/2}} \sum_x |x \otimes f(x)\rangle$$

L'ensemble des deux registres contient *simultanément* les 2^{n+m} valeurs du couple $(x, f(x))$, bien qu'une mesure donne toujours un couple unique. L'art du calcul quantique consiste à construire des transformations unitaires conduisant à des états finaux tels que les valeurs de x qui nous intéressent soient associées à des probabilités $|a_x|^2$ importantes. C'est une caractéristique des algorithmes quantiques de ne pas donner un résultat certain, mais de donner le résultat avec une certaine probabilité que l'on essaie de rendre aussi grande que possible. Un algorithme *classique* permet ensuite de vérifier rapidement que le résultat est correct : par exemple, dans le cas de l'algorithme de Shor qui permet de décomposer un nombre entier en facteurs premiers, une fois que l'algorithme quantique a proposé une solution possible pour la factorisation, cette solution peut être vérifiée rapidement sur un ordinateur classique. Nous allons illustrer la méthode sur *l'algorithme de recherche de Grover*. C'est un algorithme qui permet de rechercher une entrée dans une base de données *non structurée*, par exemple un numéro de téléphone dans un annuaire lorsque l'on connaît le numéro, mais pas la personne dont on veut trouver le nom. Si N est le nombre d'entrées dans la base, un algorithme classique doit effectuer en moyenne $N/2$ essais : il n'y a pas d'autre possibilité que d'examiner une à une toutes les entrées. L'algorithme de Grover permet de résoudre le problème en $\sim \sqrt{N}$ opérations.

Soit n le nombre minimum de qubits nécessaire pour représenter le nombre N ; afin de simplifier les notations, nous prendrons $N = 2^n$. L'algorithme de Grover va utiliser n qubits stockés dans le registre de données, plus un

26. En dehors du cas où la correspondance $x \leftrightarrow f(x)$ est bijective, la transformation $x \to f(x)$ ne peut pas être réversible : on ne peut pas avoir de transformation $U|x\rangle = |f(x)\rangle$.

qubit auxiliaire stocké dans le registre de résultats. On part de l'état (voir la figure 11.13)

$$|\Psi\rangle = (\mathrm{H}^{\otimes n}|0^{\otimes n}\rangle) \otimes \mathrm{H}|1\rangle = \left(\frac{1}{2^{n/2}}\sum_x |x\rangle\right) \otimes (|0\rangle - |1\rangle) = |\Phi\rangle \otimes (|0\rangle - |1\rangle)$$

où le dernier qubit est le qubit auxiliaire et $|\Phi\rangle$ le vecteur (11.124). On applique la porte U_f

$$U_f|\Psi\rangle = \frac{1}{2^{n/2}}\left(\sum_x (-1)^{f(x)}|x\rangle\right) \otimes (|0\rangle - |1\rangle)$$

En effet,

- Si $f(x) = 0$, $(|0\rangle - |1\rangle) \to (|0\rangle - |1\rangle)$
- Si $f(x) = 1$, $(|0\rangle - |1\rangle) \to -(|0\rangle - |1\rangle)$.

Dans cette opération, le qubit auxiliaire reste non intriqué avec les n autres qubits. Il est commode de définir un opérateur O, "l'oracle", dont l'action est la suivante dans la base de calcul

$$\mathrm{O}|x\rangle = (-1)^{f(x)}|x\rangle \tag{11.127}$$

ce qui permet de limiter le raisonnement aux n qubits du registre de données.

La base de données est stockée au moyen de n qubits et on définit la fonction $f(x)$, $x = \{0, 1, \ldots, 2^n - 1\}$ telle que $f(x) = 0$ si $x \neq y$ et $f(x) = 1$ si $x = y$ est solution : $f(x) = \delta_{xy}$. Rappelons que la valeur de y est connue, mais on ne sait pas où cette valeur se trouve dans la base[27]. Pour simplifier l'argument, on suppose que la valeur de y est unique. L'opérateur de Grover G est défini par

$$\mathrm{G} = \mathrm{H}^{\otimes n}\,\mathrm{X}\,\mathrm{H}^{\otimes n}\,\mathrm{O} = \mathrm{H}^{\otimes n}\big(2|0\rangle\langle 0| - I\big)\mathrm{H}^{\otimes n}\,\mathrm{O} \tag{11.128}$$

où

$$\mathrm{X}|x\rangle = -(-1)^{\delta_{x0}}|x\rangle = (2|0\rangle\langle 0| - I)|x\rangle$$

Pour simplifier l'écriture, nous allons nous servir du vecteur $|\Phi\rangle$ (11.124). Compte tenu de $\mathrm{H}^2 = I$, on déduit

$$\mathrm{H}^{\otimes n}\big(2|0\rangle\langle 0| - I\big)\mathrm{H}^{\otimes n} = 2\mathrm{H}^{\otimes n}|0\rangle\langle 0|\mathrm{H}^{\otimes n} - I = 2|\Phi\rangle\langle\Phi| - I$$

et donc

$$\mathrm{G} = \big(2|\Phi\rangle\langle\Phi| - I\big)\mathrm{O} \tag{11.129}$$

Cette construction permet de dessiner le circuit logique quantique correspondant à G (figure 11.13).

27. On connaît la valeur de y (le numéro de téléphone), mais pas les données associées à y dans la base (le nom de l'abonné).

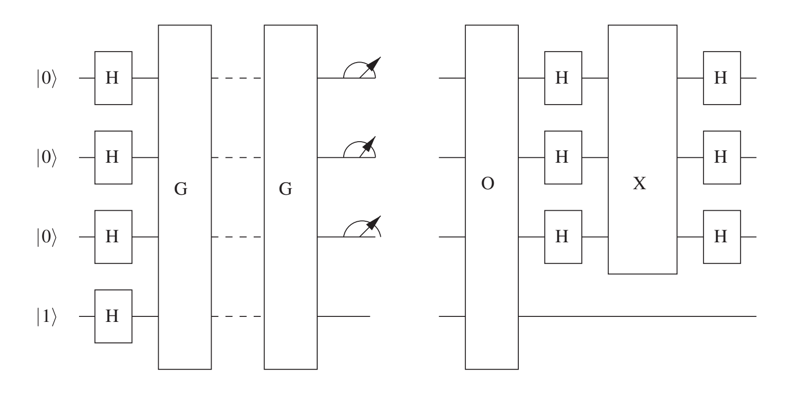


FIG. 11.13 – Circuits logiques de l'algorithme de Grover pour $n = 3$: 3 qubits et un qubit auxiliaire, qui reste non intriqué. Les circuits de G sont détaillés sur la figure de droite. L'action de l'oracle O est $\mathrm{O}|x\rangle = (-1)^{f(x)}|x\rangle$ et celle de la boîte X : $\mathrm{X}|x\rangle = -(-1)^{\delta_{x0}}|x\rangle$.

La première itération de l'algorithme va nous montrer comment le résultat souhaité est obtenu en utilisant un mécanisme d'interférences. Après application de l'opérateur de Grover sur $H^{\otimes n}|0\rangle$, le vecteur d'état du registre de données est

$$
\begin{aligned}
|\Phi'\rangle &= \frac{1}{\sqrt{N}} \sum_x (-1)^{f(x)} \big(2|\Phi\rangle\langle\Phi| - I\big)|x\rangle \\
&= \left(\sum_z \frac{2}{N}\,(-1)^{f(z)}\right)|\Phi\rangle - \frac{1}{\sqrt{N}}\sum_x (-1)^{f(x)}|x\rangle \\
&= \frac{1}{\sqrt{N}} \sum_x \left(\frac{2}{N}\,\sum_z (-1)^{f(z)} - (-1)^{f(x)}\right)|x\rangle \\
&= \sum_x a_x |x\rangle \qquad (11.130)
\end{aligned}
$$

où nous avons utilisé $\langle\Phi|x\rangle = 1/\sqrt{N}$. Prenons pour fixer les idées $n = 4$, $N = 16$. Les amplitudes $a_{x=y}$ et $a_{x\neq y}$ sont différentes ; d'après (11.130)

$$
a_{x\neq y} = \frac{3}{16} \qquad a_{x=y} = \frac{11}{16}
$$

Au départ, chaque valeur de x avait une probabilité de $1/16 \simeq 0.06$, alors que la probabilité de trouver la valeur y dans la base de données après application de G est $121/256 \simeq 0.47$, soit près de 50 %. Le raisonnement précédent peut manifestement servir à établir une relation de récurrence pour les applications successives de G, mais il est plus commode d'avoir recours à une méthode géométrique que le lecteur trouvera exposée dans les références.

Le gain de l'algorithme de Grover est limité : on passe de N opérations à $\sqrt{N}$ opérations. Un exemple où le gain est exponentiel est celui de la détermination de la période d'une fonction $f(x)$. Supposons $f(x)$ définie sur Z_N, le groupe additif des entiers modulo N ; un algorithme mis en œuvre sur un ordinateur classique exige un nombre d'opérations de l'ordre de N, alors qu'avec un ordinateur quantique, ce nombre serait $O(\ln^2 N)$. Ce résultat est à la base de l'algorithme de Shor pour la décomposition d'un nombre en facteurs premiers, la fonction $f(x)$ étant alors a^xmod. N, a entier. Pour un entier N, cet algorithme requiert $O(\ln^3 N)$ opérations sur un ordinateur quantique, alors que les meilleurs algorithmes classiques exigent $O(\exp[(\ln N)^{1/3}])$ opérations.

Le principe d'algorithmes fonctionnant sur des ordinateurs quantiques étant acquis, reste la réalisation concrète d'un tel ordinateur. Les avis sur cette question sont partagés : ils vont du pessimisme intégral à un optimisme mesuré. À l'heure actuelle, le record en nombre de qubits est détenu par un groupe d'IBM (Vandersypen *et al.* [2001]), qui a obtenu la factorisation de 15 : $15 = 3 \times 5$ (!) à l'aide d'un ordinateur quantique utilisant la RMN, mais on est encore très loin de résultats utiles. Un autre schéma prometteur est fondé sur la manipulation d'ions piégés (exercice 11.6.15). Le principal problème est celui de la décohérence : en effet, les algorithmes quantiques exigent que l'évolution soit unitaire, ce qui implique l'absence d'interactions incontrôlées avec l'environnement, Bien sûr, isoler complètement l'ordinateur quantique est impossible : il s'agit de réduire au maximum les perturbations induites par l'environnement et de concevoir des algorithmes de correction d'erreurs inévitables en utilisant une information redondante. En dépit des difficultés, le domaine du calcul quantique est en pleine expansion.

11.5.3 Téléportation quantique

La *téléportation quantique* permet de transférer de l'information quantique d'un endroit à un autre sans que ce transfert s'accompagne de la propagation entre ces deux endroits d'une particule qui porte l'information. Elle pourrait avoir des applications pour la cryptographie quantique (figure 11.14). Supposons qu'Alice souhaite transférer à Bob l'information sur l'état de spin $|\varphi_A\rangle$ d'une particule $\mathcal{A}$ de spin 1/2

$$|\varphi_A\rangle = \lambda|0_A\rangle + \mu|1_A\rangle \tag{11.131}$$

qui lui est *a priori* inconnu, sans lui transmettre directement cette particule. Elle ne peut pas faire une mesure du spin, car elle ne connaît pas l'orientation du spin de la particule $\mathcal{A}$, et toute mesure projetterait en général $|\varphi_A\rangle$ sur un autre état. Le principe du transfert de l'information consiste à utiliser une paire auxiliaire de particules intriquées $\mathcal{B}$ et $\mathcal{C}$ de spin 1/2 partagées par Alice et Bob : la particule $\mathcal{B}$ est utilisée par Alice et la particule $\mathcal{C}$ est envoyée vers Bob (figure 11.14). Ces particules $\mathcal{B}$ et $\mathcal{C}$ se trouvent, par exemple, dans l'état

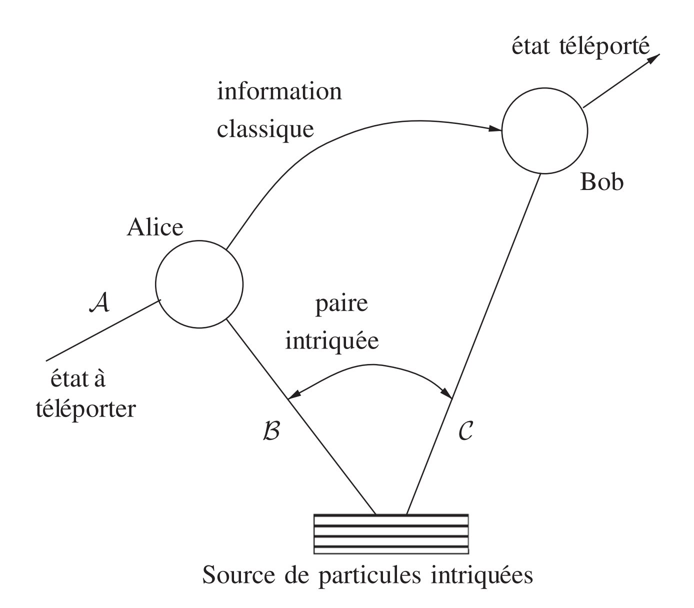


FIG. 11.14 – Téléportation : Alice effectue une mesure de Bell sur les qubits A et B et informe Bob du résultat par une voie classique.

intriqué de spin $|\Phi^{(+)}_{BC}\rangle$ (voir (11.137))

$$|\Phi^{(+)}_{BC}\rangle = \frac{1}{\sqrt{2}}\big(|0_B 0_C\rangle + |1_B 1_C\rangle\big) \tag{11.132}$$

L'état initial des trois particules $|\Phi_{ABC}\rangle$ est donc

$$\begin{aligned}|\Phi_{ABC}\rangle &= \big(\lambda|0_A\rangle + \mu|1_A\rangle\big)\frac{1}{\sqrt{2}}\big(|0_B 0_C\rangle + |1_B 1_C\rangle\big)\\ &= \frac{\lambda}{\sqrt{2}}\,|0_A\rangle\big(|0_B 0_C\rangle + |1_B 1_C\rangle\big) + \frac{\mu}{\sqrt{2}}\,|1_A\rangle\big(|0_B 0_C\rangle + |1_B 1_C\rangle\big)\end{aligned} \tag{11.133}$$

Alice va d'abord appliquer sur les qubits $\mathcal{A}$ et $\mathcal{C}$ une porte cNOT, le qubit $\mathcal{A}$ jouant le rôle de qubit de contrôle et le qubit $\mathcal{B}$ celui de qubit cible (figure 11.15). Cette opération transforme l'état initial (11.133) à trois qubits en

$$|\Phi'_{ABC}\rangle = \frac{\lambda}{\sqrt{2}}\big(|0_A\rangle(|0_B 0_C\rangle + |1_B 1_C\rangle\big) + \frac{\mu}{\sqrt{2}}\big(|1_A\rangle(|1_B 0_C\rangle + |0_B 1_C\rangle\big) \tag{11.134}$$

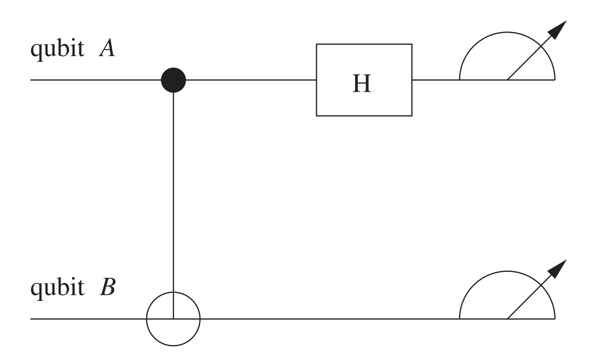


FIG. 11.15 – Alice applique une porte cNOT sur les qubits $\mathcal{A}$ et $\mathcal{B}$ puis une porte de Hadamard sur le qubit $\mathcal{A}$.

Alice applique ensuite une porte de Hadamard sur le qubit $\mathcal{A}$, ce qui transforme (11.134) en

$$|\Phi''_{ABC}\rangle = \frac{1}{2}\Big[\lambda|0_A0_B0_C\rangle + \lambda|0_A1_B1_C\rangle + \lambda|1_A0_B0_C\rangle + \lambda|1_A1_B1_C\rangle + \mu|0_A1_B0_C\rangle + \mu|0_A0_B1_C\rangle - \mu|1_A1_B0_C\rangle - \mu|1_A0_B1_C\rangle\Big] \quad (11.135)$$

Cette équation peut se récrire

$$|\Phi''_{ABC}\rangle = \frac{1}{2}\,|0_A0_B\rangle\big(\lambda|0_C\rangle + \mu|1_C\rangle\big) + \frac{1}{2}\,|0_A1_B\rangle\big(\mu|0_C\rangle + \lambda|1_C\rangle\big) + \frac{1}{2}\,|1_A0_B\rangle\big(\lambda|0_C\rangle - \mu|1_C\rangle\big) + \frac{1}{2}\,|1_A1_B\rangle\big(-\mu|0_C\rangle + \lambda|1_C\rangle\big) \quad (11.136)$$

La dernière opération d'Alice consiste à mesurer les deux qubits dans la base $\{|0\rangle, |1\rangle\}$. La mesure conjointe par Alice des qubits $\mathcal{A}$ et $\mathcal{B}$ est appelée *mesure de Bell*. Cette mesure projette la paire $\mathcal{AB}$ sur l'un des quatre états $|i_Aj_B\rangle$, $i, j = 0, 1$, et le vecteur d'état du qubit $\mathcal{C}$ se lit sur chacune des lignes de (11.136).

Le cas le plus simple est celui où le résultat de la mesure est $|0_A0_B\rangle$. Le qubit $\mathcal{C}$ arrive alors à Bob dans l'état

$$\lambda|0_C\rangle + \mu|1_C\rangle$$

c'est-à-dire dans l'état initial du qubit $\mathcal{A}$, avec les *mêmes* coefficients (complexes !) λ et μ. Alice informe Bob par une voie classique (téléphone...) que le qubit va lui arriver dans le même état que le qubit $\mathcal{A}$. Si au contraire, elle mesure $|0_A1_B\rangle$, le qubit $\mathcal{C}$ est dans l'état

$$\mu|0_C\rangle + \lambda|1_C\rangle$$

et elle informe Bob qu'il doit appliquer au qubit $\mathcal{C}$ une rotation de π autour de Ox, ou de façon équivalente la matrice σ_x

$$\exp\left(-\mathrm{i}\frac{\pi\sigma_x}{2}\right) = -\mathrm{i}\sigma_x$$

Dans le troisième cas ($|1_A 0_B\rangle$), il faut appliquer une rotation de π autour de Oz, et dans le dernier cas ($|1_A 1_B\rangle$) une rotation de π autour de Oy. On note que dans les quatre cas de figure, Alice ne connaît pas les coefficients λ et μ, et elle communique uniquement à Bob les informations sur la rotation qu'il doit effectuer.

Il est utile d'ajouter les remarques finales :

- à aucun moment les coefficients λ et μ ne sont mesurés, et l'état $|\varphi_A\rangle$ est détruit au cours de la mesure faite par Alice. Il n'y a donc pas de contradiction avec le théorème de non-clonage ;
- Bob ne "connaît" l'état de la particule $\mathcal{C}$ que lorsqu'il a reçu le résultat de la mesure d'Alice. La transmission de cette information doit se faire par une voie classique, à une vitesse au plus égale à celle de la lumière. Il n'y a donc pas transmission instantanée de l'information à distance ;
- il n'y a jamais transport de matière dans la téléportation quantique[28].

11.5.4 Échange d'intrication

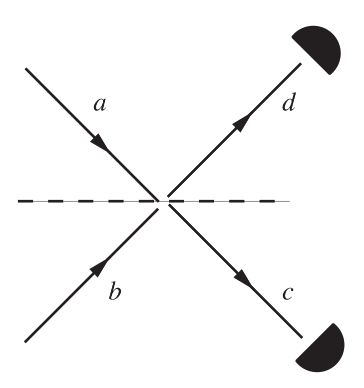


FIG. 11.16 – Deux photons arrivant sur une lame séparatrice.

Jusqu'à présent, nous avons décrit des états intriqués produits par une source (atome, cristal non linéaire...), et les deux particules de l'état intriqué ont été en contact ou très proches dans le passé. En fait, grâce à l'échange d'intrication, on peut parfaitement obtenir des états intriqués de deux particules qui n'ont jamais été en contact, n'ont jamais eu de passé commun,

28. Pour une expérience récente de téléportation, voir par exemple Olmschenk *et al.* [2009], qui utilisent le schéma décrit dans le § 11.5.4.

et peuvent même avoir toujours été arbitrairement éloignées. Nous allons décrire une expérience d'échange d'intrication effectivement réalisée (Olmschenk *et al.* [2009]). Commençons par montrer un résultat préliminaire : considérons une lame séparatrice équilibrée sur laquelle arrivent deux photons A et B, qui se trouvent dans des modes de propagation, ou modes spatiaux, a et b et sortent de la lame suivant les modes de propagation c et d (figure 11.16). Ces photons possèdent, en outre, un mode interne à deux niveaux : polarisation, couleur, etc. Un point crucial pour la suite est que le dispositif expérimental, par exemple celui qui sera décrit ci-dessous, ne permet pas *par principe* de savoir si le photon A par exemple arrive en dessous de la lame ou au-dessus. Pour décrire les modes internes, nous utilisons les états de Bell

$$
\begin{aligned}
|\Psi^{(\pm)}_{AB}\rangle &= \frac{1}{\sqrt{2}}(|0_A 1_B\rangle \pm |1_B 0_A\rangle) \\
|\Phi^{(\pm)}_{AB}\rangle &= \frac{1}{\sqrt{2}}(|0_A 0_B\rangle \pm |1_A 1_B\rangle)
\end{aligned}
\tag{11.137}
$$

La remarque capitale pour la suite est que $|\Psi^{(+)}_{AB}\rangle$ et $|\Phi^{(\pm)}_{AB}\rangle$ sont symétriques dans l'échange $A \leftrightarrow B$, alors que $|\Psi^{(-)}_{AB}\rangle$ est antisymétrique : si l'on identifie le mode interne à un spin 1/2, $|\Psi^{(+)}_{AB}\rangle$ et $|\Phi^{(\pm)}_{AB}\rangle$ correspondent à un état de spin 1 (triplet) symétrique dans l'échange des deux spins, et $|\Psi^{(-)}_{AB}\rangle$ à un état de spin 0 (singulet) antisymétrique.

Le vecteur d'état décrivant la partie spatiale peut également être symétrique ou antisymétrique. Afin d'alléger les notations et d'éviter une confusion possible, nous identifierons les photons par un indice (1,2), plutôt que par (A, B). Les vecteurs d'état symétrique (S) ou antisymétrique (A) sont

$$
\begin{aligned}
|\Psi^{S}_{12}\rangle &= \frac{1}{\sqrt{2}}(|a_1 b_2\rangle + |a_2 b_1\rangle) \\
|\Psi^{A}_{12}\rangle &= \frac{1}{\sqrt{2}}(|a_1 b_2\rangle - |a_2 b_1\rangle)
\end{aligned}
\tag{11.138}
$$

où $|a_1\rangle(|b_2\rangle)$, par exemple, indique que le photon 1 est dans le mode a (le photon 2 dans le mode b). Nous verrons au chapitre 14 que le vecteur d'état global de deux photons doit être symétrique dans l'échange des deux photons, car les photons obéissent à la statistique de Bose-Einstein. Les vecteurs d'état possibles sont donc

$$
|\Psi^{S}_{12}\rangle \otimes |\Psi^{(+)}_{12}\rangle, \quad |\Psi^{S}_{12}\rangle \otimes |\Phi^{(+)}_{12}\rangle, \quad |\Psi^{S}_{12}\rangle \otimes |\Phi^{(-)}_{12}\rangle, \quad |\Psi^{A}_{12}\rangle \otimes |\Psi^{(-)}_{12}\rangle
$$

On constate que l'état spatial $|\Psi^{A}_{12}\rangle$ est associé de façon biunivoque à l'état de Bell $|\Psi^{(-)}_{12}\rangle$.

Écrivons maintenant l'action de la lame séparatrice sur les modes spatiaux

$$|a\rangle \implies \frac{1}{\sqrt{2}}(|c\rangle + \mathrm{i}|d\rangle)$$
$$|b\rangle \implies \frac{1}{\sqrt{2}}(\mathrm{i}|c\rangle + |d\rangle)$$

On en déduit pour un état à deux photons

$$|a_1 b_2\rangle \implies \frac{1}{2}(\mathrm{i}|c_1 c_2\rangle + \mathrm{i}|d_1 d_2\rangle + |c_1 d_2\rangle - |d_1 c_2\rangle)$$
$$|a_2 b_1\rangle \implies \frac{1}{2}(\mathrm{i}|c_1 c_2\rangle + \mathrm{i}|d_1 d_2\rangle - |c_1 d_2\rangle + |d_1 c_2\rangle)$$

Un terme tel que $|c_1 c_2\rangle$ représente deux photons se propageant le long de la même voie de sortie de la lame, alors qu'un terme comme $|c_1 d_2\rangle$ représente des photons empruntant des voies de sortie différentes : dans le premier cas, un seul détecteur est déclenché, dans le second les deux détecteurs sont déclenchés simultanément. En d'autres termes, si l'état spatial des deux photons est symétrique,

$$\frac{1}{\sqrt{2}}(|c_1 c_2\rangle + |d_1 d_2\rangle)$$

un seul détecteur est déclenché, s'il est antisymétrique

$$\frac{1}{\sqrt{2}}(|c_1 d_2\rangle - |d_1 c_2\rangle)$$

les deux détecteurs sont déclenchés simultanément. Compte tenu de la correspondance biunivoque entre l'état spatial antisymétrique et l'état de Bell $|\Psi_{12}^{(-)}\rangle$, on en déduit que le déclenchement simultané des deux détecteurs implique que l'état interne est cet état de Bell. Finalement, nous montrerons au § 17.2.2 que si les deux photons arrivent sur la lame dans le même mode interne, alors ils doivent déclencher le même détecteur : c'est à nouveau une conséquence de la statistique de Bose-Einstein pour les photons.

Passons maintenant à la description de l'expérience de Olmschenk *et al.* [2009]. Cette expérience utilise des niveaux d'énergie d'ions piégés d'Ytterbium, Yb^+. Le schéma des niveaux utilisé est donné sur la figure 11.17 : les transitions se font de deux niveaux $|g\rangle$ et $|e\rangle$ vers les niveaux $|g'\rangle$ et $|e'\rangle$, respectivement. La transition $|g'\rangle \to |g\rangle$ donne un photon de longueur d'onde plus courte que celui de la transition $|e'\rangle \to |e\rangle$. Ces deux photons seront appelé conventionnellement photon "bleu" $|B\rangle$ et photon "rouge" $|R\rangle$[29]. On

29. Il s'agit, bien sûr, d'une convention de langage : la longueur d'onde moyenne de la transition est de 369.5 nm et la différence entre les deux transitions correspond à une fréquence de 14.7 MHz. Les niveaux $|g\rangle$ et $|e\rangle$ correspondent à des niveaux de structure hyperfine (§ 15.2.4) $F = 0$ et $F = 1, m_F = 0$, respectivement, tandis que les niveaux $|g'\rangle$ et $|e'\rangle$ correspondent à $F = 1, m_F = 0$ et $F = 0$, respectivement.

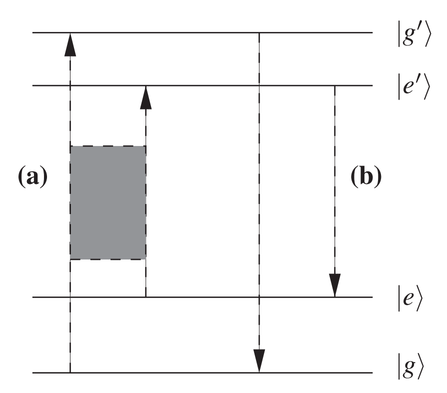


FIG. 11.17 – Schéma des niveaux. En (a), une impulsion laser à large bande fait passer des niveaux $|g\rangle$ et $|e\rangle$ aux niveaux $|e'\rangle$ et $|g'\rangle$. En (b), ces deux niveaux se désexcitent par émission spontanée. La transition $|g'\rangle \to |g\rangle$ correspond au photon "bleu", et $|e'\rangle \to |e\rangle$ au photon "rouge".

commence par préparer un ion dans une combinaison linéaire des états $|g\rangle$ et $|e\rangle$

$$|\psi\rangle = \alpha|g\rangle + \beta|e\rangle$$

Une impulsion laser à large bande fait passer simultanément de $|g\rangle$ et $|e\rangle$ à $|g'\rangle$ et $|e'\rangle$, qui ensuite émettent spontanément un photon $B(|g'\rangle)$ ou $R(|e'\rangle)$. Nous obtenons donc un état intriqué entre les niveaux de l'ion et la couleur du photon

$$|\Psi\rangle = \alpha|gB\rangle + \beta|eR\rangle \tag{11.139}$$

Cette opération peut être réalisée sur deux ions identiques et distants d'environ 1 m (figure 11.18), et l'état global ions/photons est un simple produit tensoriel

$$|\Psi_{12}\rangle = (\alpha|g_1B_1\rangle + \beta|e_1R_1\rangle) \otimes (\gamma|g_2B_2\rangle + \delta|e_2R_2\rangle \tag{11.140}$$

Les photons émis par les deux ions sont ensuite recueillis par une lentille et envoyés sur une lame séparatrice ; les deux modes a et b précédents sont les couleurs B et R. Comme nous l'avons mentionné, si les deux photons sont de la même couleur, ils partent vers le même détecteur. Si les deux détecteurs sont déclenchés simultanément, cela veut dire que les deux photons sont de couleur différente et il suffit de prendre en compte dans (11.140) la partie

$$|\Psi'\rangle = \alpha\delta|g_1e_2\rangle \otimes |B_1R_2\rangle + \beta\gamma|e_1g_2\rangle \otimes |R_1B_2\rangle \tag{11.141}$$

Exprimons $|\Psi'\rangle$ en fonction des états de Bell (11.137) (*cf.* l'exercice 11.6.10 pour une étude plus détaillée)

$$|\Psi_{12}^{(\pm)}\rangle = \frac{1}{\sqrt{2}}(|B_1R_2\rangle \pm |R_1B_2\rangle) \tag{11.142}$$

ce qui donne

$$
\begin{aligned}
|\Psi'\rangle = & \frac{1}{\sqrt{2}}(\alpha\delta|g_1e_2\rangle + \beta\gamma|e_1g_2\rangle) \otimes |\Psi_{12}^{(+)}\rangle \\
& + \frac{1}{\sqrt{2}}(\alpha\delta|g_1e_2\rangle - \beta\gamma|e_1g_2\rangle) \otimes |\Psi_{12}^{(-)}\rangle
\end{aligned}
\tag{11.143}
$$

De la discussion précédente, il résulte que le déclenchement simultané des deux détecteurs projette le vecteur d'état de couleur sur $|\Psi_{12}^{(-)}\rangle$, et le vecteur d'état des deux ions est par conséquent

$$
|\Phi_{12}\rangle = \alpha\delta|g_1e_2\rangle - \beta\gamma|e_1g_2\rangle \tag{11.144}
$$

Les états des ions 1 et 2 sont intriqués alors que les deux ions n'ont jamais interagi ! Bien évidemment l'intrication ne réussit pas à tous les coups, loin de là. En premier lieu, 75 % des photons sont éliminés par un filtrage en polarisation, ensuite seulement 1 % des photons émis sont collectés par les lentilles, les détecteurs ont une efficacité quantique de 15 %, etc., et au bout du compte, pour chaque essai, la probabilité de réussite est $\sim 10^{-8}$. Mais l'observation du déclenchement simultané des deux détecteurs garantit que le processus a réussi.

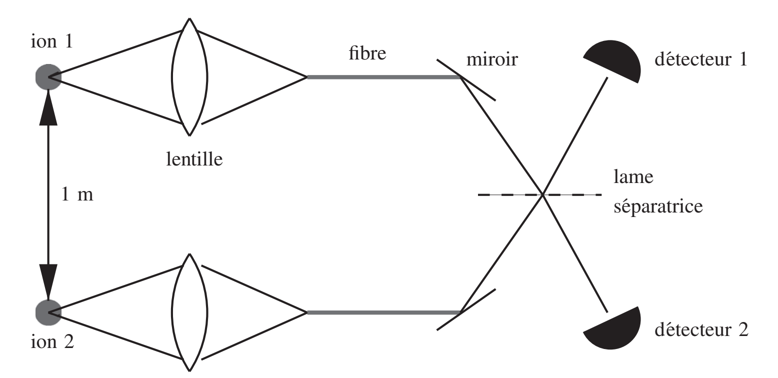


FIG. 11.18 – Intrication de deux ions par détection de photons émis.

Le schéma général d'échange d'intrication est le suivant (figure 11.19) : deux couples de particules AB et CD sont produites dans un état intriqué, par exemple un état de Bell $|\Psi^{(+)}\rangle$ (11.137)

$$
\begin{aligned}
|\Psi_{AB}^{(+)}\rangle &= \frac{1}{\sqrt{2}}(|0_A1_B\rangle + |1_A0_B\rangle) \\
|\Psi_{CD}^{(+)}\rangle &= \frac{1}{\sqrt{2}}(|0_C1_D\rangle + |1_C0_D\rangle)
\end{aligned}
\tag{11.145}
$$

L'état initial des 4 particules est un produit tensoriel $|\Psi\rangle = |\Psi_{AB}^{(+)}\rangle \otimes |\Psi_{CD}^{(+)}\rangle$ qui peut s'écrire (exercice 11.6.10) en fonction d'états de Bell formés avec les

paires AD et BC

$$|\Psi\rangle = \frac{1}{2}\Big(|\Psi^{(+)}_{AD}\rangle \otimes |\Psi^{(+)}_{BC}\rangle - |\Psi^{(-)}_{AD}\rangle \otimes |\Psi^{(-)}_{BC}\rangle + |\Phi^{(+)}_{AD}\rangle \otimes |\Phi^{(+)}_{BC}\rangle - |\Phi^{(-)}_{AD}\rangle \otimes |\Phi^{(-)}_{BC}\rangle\Big) \qquad (11.146)$$

Si une mesure de Bell effectuée sur la paire BC a pour résultat un des 4 états de Bell, alors la paire AD est envoyée dans l'état de Bell correspondant.

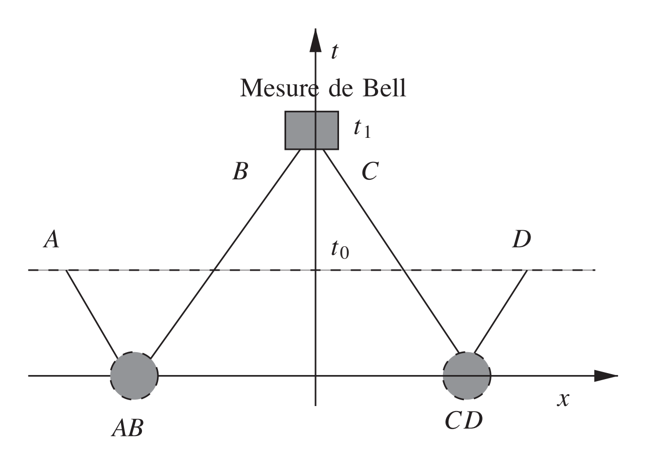


FIG. 11.19 – Échange d'intrication. Deux sources produisent deux paires intriquéees AB et CD. Une mesure de Bell faite sur la paire BC entraîne l'intrication de A et D. La mesure de (BC) est effectuée longtemps après celle de A et D ($t_0 \ll t_1$).

Sur le schéma de la figure 11.19, les instants où sont effectuées les mesures de polarisation sont totalement arbitraires. Pour fixer les idées, prenons pour particules intriquées des photons dont la polarisation est 0 si elle est parallèle à l'axe de mesure et 1 si elle est perpendiculaire à cet axe. On peut imaginer une situation dite de *post-sélection* où, dans un premier temps, on accumule des données sur la polarisation d'un grand nombre de photons A et D, et *une fois toutes ces données enregistrées*, on mesure la polarisation des photons B et C. On peut, par exemple, mesurer les polarisations (BC) à une grande distance des sources ou envoyer les photons dans plusieurs dizaines de km de fibres optiques. Les résultats enregistrés pour les polarisations des photons A et D sont des suites aléatoires de 0 et de 1, et nous allons montrer que ces suites aléatoires enregistrées une fois pour toutes peuvent être interprétées de deux façons radicalement différentes, *après que la mesure (BC) a été effectuée.*

Supposons d'abord que les polarisations des photons B et C soient mesurées indépendamment. La série des valeurs observées pour le photon A par exemple est une suite aléatoire de 0 et de 1, mais ces valeurs sont (anti-)corrélées à celles observées pour le photon B : si les résultats pour B sont $\{011001\}$, alors les résultats pour A suivant le même axe sont $\{100110\}$

en raison de (11.145). En utilisant pour chaque photon des axes d'orientations différentes, on pourra même vérifier la borne de Cirelson et en déduire que les photons A et B étaient dans un état intriqué $|\Psi^{(+)}_{AB}\rangle$. Le même argument est valable pour la paire (CD).

Cependant, supposons maintenant que la mesure de la paire (BC) soit faite dans la base de Bell (11.137). Lorsque la mesure est effectuée, il y a une chance sur 4 pour que la paire soit trouvée dans un des états de Bell (11.137), par exemple dans l'état de Bell $|\Psi^{(+)}_{BC}\rangle$, et on peut trier les résultats (BC) selon le type d'état de Bell enregistré. D'après (11.146), toutes les paires (AD) correspondant à un résultat $|\Psi^{(+)}_{BC}\rangle$ seront alors intriquées dans l'état $|\Psi^{(+)}_{AD}\rangle$, et la comparaison des mesures de polarisation sur A et D permettra de le vérifier. Autrement dit, les mesures des polarisations des photons A et D, qui étaient *a priori* des suites aléatoires de 0 et de 1, prennent une signification précise, et *différente* selon le type de mesure effectué sur les deux autres photons, *postérieurement* à l'enregistrement des polarisations de la paire (AD) : c'est ce que l'on appelle une procédure de post-sélection. La post-sélection existe aussi pour les probabilités classiques. Ce qui est spécifique de l'intrication quantique est la violation des inégalités de Bell. Le schéma de la figure 11.19 est à la base des relais quantiques susceptibles de transporter l'information quantique sur de longues distances.

11.6 Exercices

11.6.1 Propriétés des opérateurs statistiques

1. On construit avec les éléments de matrice ρ_{ii}, ρ_{ij}, ρ_{ji} et ρ_{jj} d'un opérateur statistique ρ dans un espace de dimension d la matrice 2×2

$$A = \begin{pmatrix} \rho_{ii} & \rho_{ij} \\ \rho_{ji} & \rho_{jj} \end{pmatrix}$$

Montrer que $\rho_{ii} \geq 0$, $\rho_{jj} \geq 0$ et que $\det A \geq 0$, d'où $|\rho_{ij}|^2 \leq \rho_{ii}\rho_{jj}$. En déduire également que si $\rho_{ii} = 0$, alors $\rho_{ij} = \rho^*_{ji} = 0$.

2. Montrer que s'il existe un test maximal donnant une probabilité de 100 % pour l'état physique décrit par un opérateur statistique ρ, alors cet état est un cas pur. Montrer également que si ρ décrit un cas pur, et que l'on peut écrire

$$\rho = \lambda\rho' + (1-\lambda)\rho'' \qquad 0 \leq \lambda \leq 1$$

alors $\rho = \rho' = \rho''$.

3. Montrer que $\mathrm{Tr}\rho^2 = 1$ est une condition nécessaire et suffisante pour un cas pur.

4. Vérifier que ρ défini par (11.8) obéit bien à toutes les propriétés d'un opérateur statistique.

11.6.2 Structure fine et effet Zeeman du positronium

Le positronium est un état lié électron-positron très semblable à l'état lié électron-proton de l'atome d'hydrogène.

1. Calculer l'énergie de l'état fondamental du positronium en fonction de celui de l'atome d'hydrogène. On rappelle que la masse du positron est égale à celle de l'électron.

2. Dans la suite de l'exercice, on s'intéressera uniquement à la structure en spin de l'état fondamental du positronium. L'espace des états à prendre en compte est donc un espace $\mathcal{H}$ à quatre dimensions, produit tensoriel des espaces des états de spin 1/2 de l'électron et du positron. Suivant les notations du § 11.1.2, on notera $|\varepsilon_1\varepsilon_2\rangle$ un état où la composante z du spin de l'électron est $\hbar\varepsilon_1/2$ et celle du positron $\hbar\varepsilon_2/2$, avec $\varepsilon = \pm 1$. Déterminer l'action des opérateurs $\sigma_{1x}\sigma_{2x}$, $\sigma_{1y}\sigma_{2y}$ et $\sigma_{1z}\sigma_{2z}$ sur les quatre états de base $|++\rangle$, $|+-\rangle$, $|-+\rangle$ et $|--\rangle$ de $\mathcal{H}$. En déduire l'action sur ces états de

$$\vec{\sigma}_1 \cdot \vec{\sigma}_2 = \sigma_{1x}\sigma_{2x} + \sigma_{1y}\sigma_{2y} + \sigma_{1z}\sigma_{2z}$$

3. Montrer que les quatre vecteurs

$$\begin{aligned} |I\rangle &= |++\rangle \\ |II\rangle &= \frac{1}{\sqrt{2}}\left(|+-\rangle + |-+\rangle\right) \\ |III\rangle &= |--\rangle \\ |IV\rangle &= \frac{1}{\sqrt{2}}\left(|+-\rangle - |-+\rangle\right) \end{aligned}$$

forment une base orthonormée de $\mathcal{H}$ et que ces vecteurs sont vecteurs propres de $\vec{\sigma}_1 \cdot \vec{\sigma}_2$ avec des valeurs propres 1 et -3. Que représentent ces états ? Faire le lien avec le § 9.6.1.

4. Déterminer les projecteurs $\mathcal{P}_1$ et $\mathcal{P}_{-3}$ sur les sous-espaces des valeurs propres 1 et -3, en écrivant ces projecteurs sous la forme

$$\lambda I + \mu \vec{\sigma}_1 \cdot \vec{\sigma}_2$$

5. Montrer que l'opérateur $\mathcal{P}_{12}$

$$\mathcal{P}_{12} = \frac{1}{2}\left(I + \vec{\sigma}_1 \cdot \vec{\sigma}_2\right)$$

échange les valeurs de ε_1 et ε_2

$$\mathcal{P}_{12}|\varepsilon_1\varepsilon_2\rangle = |\varepsilon_2\varepsilon_1\rangle$$

6. Le hamiltonien H_0 du système de spins est donné en l'absence de champ extérieur par

$$H_0 = E_0 I + A\vec{\sigma}_1 \cdot \vec{\sigma}_2 \qquad A > 0$$

où E_0 et A sont des constantes. Déterminer les vecteurs propres et valeurs propres de H_0.

7. Le positronium est placé dans un champ magnétique $\vec{B}$ uniforme et constant parallèle à Oz. Montrer que le hamiltonien devient

$$H = H_0 - \frac{q_e \hbar}{2m} B(\sigma_{1z} - \sigma_{2z})$$

où m est la masse de l'electron et q_e sa charge. Déterminer la matrice représentative de H dans la base $\{|I\rangle, |II\rangle, |III\rangle, |IV\rangle\}$. On définit le paramètre x par

$$\frac{q_e \hbar}{2m} B = -Ax$$

Déterminer les valeurs propres de H et tracer sur un graphique leur dépendance en fonction de x.

11.6.3 Ondes de spin et magnons

NB : Cet exercice utilise les notations et les résultats des questions **2** à **5** de l'exercice précédent. On peut représenter un corps ferromagnétique à une dimension comme une chaîne de spins 1/2 : N spins 1/2 numérotés $n = 0, \ldots, N-1$, $N \gg 1$, qui sont disposés en chaque point d'un réseau à une distance l l'un de l'autre. Il sera commode d'utiliser des conditions aux limites périodiques, où le spin N est identifié au spin 0 : $N \equiv 0$. On suppose que chaque spin peut interagir uniquement avec ses deux plus proches voisins et le hamiltonien s'écrit en fonction d'une constante A

$$H = \frac{1}{2} NAI - \frac{1}{2} A \sum_{n=0}^{N-1} \vec{\sigma}_n \cdot \vec{\sigma}_{n+1}$$

1. Montrer que toute valeur propre E de H vérifie $E \geq 0$ et que le minimum E_0 correspondant à l'état fondamental est atteint quand tous les spins sont orientés dans la même direction. Dans la suite de l'exercice, on choisira cette direction comme axe des z. Un choix possible pour l'état fondamental $|\Phi_0\rangle$ est alors[30]

$$|\Phi_0\rangle = |+++\ldots+++\rangle$$

2. Montrer que H s'écrit

$$H = NAI - A \sum_{n=0}^{N-1} \mathcal{P}_{n,n+1} = A \sum_{n=0}^{N-1} (I - \mathcal{P}_{n,n+1})$$

où

$$\mathcal{P}_{n,n+1} = \frac{1}{2} (I + \vec{\sigma}_n \cdot \vec{\sigma}_{n+1})$$

30. Tout état obtenu à partir de $|\Phi_0\rangle$ par une rotation de l'ensemble des spins d'un même angle autour d'un même axe est encore un état fondamental possible.

En utilisant le résultat de la question **5** de l'exercice précédent, montrer que les vecteurs propres de H sont des combinaisons linéaires de vecteurs où le nombre de spins up moins le nombre de spins down est une constante. Soit $|\Psi_n\rangle$, l'état où le spin n est down, tous les autres spins étant up. Quelle est l'action de H sur $|\Psi_n\rangle$?

3. On cherche des vecteurs propres $|k_s\rangle$ de H comme combinaisons linéaires des $|\Psi_n\rangle$. Compte tenu de la symétrie cyclique, on pose

$$|k_s\rangle = \sum_{n=0}^{N-1} \mathrm{e}^{\,\mathrm{i}k_s nl}|\Psi_n\rangle$$

avec

$$k_s = \frac{2\pi s}{Nl} \qquad s = 0,\ 1,\ \ldots\ N-1$$

Montrer que $|k_s\rangle$ est vecteur propre de H et déterminer l'énergie correspondante E_k. Montrer que l'énergie est proportionnelle à k_s^2 si $k_s \to 0$. On associe à l'état $|k_s\rangle$ de (pseudo-)vecteur d'onde k_s et d'énergie E_k une particule appelée *magnon*.

11.6.4 Écho de spin et décomposition des niveaux en RMN

1. Pour diverses raisons, il est important de mesurer avec précision le temps de relaxation transverse T_2 (§ 5.2.3 et § 11.4.3) dans des expériences de RMN. Dans le référentiel en rotation, le signal RMN $a(t)$ prend la forme (δ est le désaccord)

$$a(t) \propto \mathrm{e}^{\mathrm{i}\delta t/2}\, \mathrm{e}^{-t/T_2}.$$

Calculer la transformée de Fourier $\tilde{a}(\omega)$ de $a(t)$

$$\tilde{a}(\omega) = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t}\, a(t).$$

On pourrait espérer déduire T_2 de la largeur $1/T_2$ du pic de la partie réelle de $\tilde{a}(\omega)$. Cependant, les différentes molécules auront des désaccords différents, par exemple parce que $\vec{B}_0$ peut être légèrement inhomogène, ce qui conduit à des fréquences de Larmor différentes, de sorte que les signaux des différentes molécules interfèrent destructivement, et $a(t)$ décroît avec un temps caractéristique beaucoup plus petit que T_2. Pour contourner cette difficulté, on applique la séquence suivante d'opérations sur la matrice statistique (11.39) : évolution libre pendant $t/2$, rotation de π autour de l'axe y et évolution libre pendant $t/2$. Montrer qu'en l'absence de relaxation, la matrice statistique évoluerait à partir de $\rho(t=0)$ (11.39) selon

$$\begin{aligned} \rho(t=0) \to \rho(t) &= U(t)\,\rho(t=0)\,U^\dagger(t) \\ U(t) &= \exp\left(\frac{-\mathrm{i}\delta\sigma_z t}{4}\right)(-\mathrm{i}\sigma_y)\exp\left(\frac{-\mathrm{i}\delta\sigma_z t}{4}\right) \end{aligned}$$

Montrer que $U(t) = -\mathrm{i}\sigma_y$, et que si l'on prend en compte la relaxation $\rho(t)$ est

$$\rho(t) = \frac{1}{2}\left(I + \frac{1}{2}\,\delta\mathsf{p}\sigma_y\,\mathrm{e}^{-t/T_2}\right)$$

indépendamment du désaccord δ. Montrer que la mesure de la décroissance de la hauteur du pic de $\tilde{a}(\omega)$ permet une détermination fiable de T_2, et expliquer pourquoi cette séquence d'opérations s'appelle “écho de spin”.

2. Considérons deux noyaux identiques de spin 1/2 (par exemple deux protons) appartenant à une même molécule observée dans un expérience de RMN. Les deux spins nucléaires ont un hamiltonien d'interaction H_{12}, qui, dans le cas le plus simple, a la forme suivante

$$H_{12} = \hbar\omega_{12}\,\sigma_z^{(1)} \otimes \sigma_z^{(2)}$$

Montrer que l'opérateur d'évolution correspondant est donné par

$$U_{12}(t) = \exp(-\mathrm{i}H_{12}t/\hbar) = I_{12}\cos\omega_{12}t - \mathrm{i}[\sigma_z^{(1)} \otimes \sigma_z^{(2)}]\sin\omega_{12}t$$

Montrer l'identité suivante

$$U[\mathcal{R}_x^{(1)}(\pi)]\exp(-\mathrm{i}H_{12}t/\hbar)U[\mathcal{R}_x^{(1)}(\pi)]\exp(-\mathrm{i}H_{12}t/\hbar) = I_{12}$$

où $U[\mathcal{R}_x^{(1)}(\pi)]$ est une rotation de π du spin 1 autour de l'axe x. À partir de cette équation, montrer que la séquence d'opérations

$$\begin{aligned}&\text{évolution pendant } t \to \text{ rotation de } \pi \text{ autour de } Ox\\ &\to \text{évolution pendant } t \to \text{rotation de } \pi \text{ autour de } Ox\end{aligned}$$

ramène les spins à leur configuration originale à $t = 0$. La séquence précédente d'opérations est utilisée en calcul quantique RMN. Elle repose sur la propriété suivante : ω_{12}^{-1} est de l'ordre de la centaine de millisecondes, tandis qu'une rotation prend quelques dizaines de microsecondes.

3. Montrer que le hamiltonien complet des deux spins est, dans le référentiel en rotation

$$H_{\text{tot}} = \frac{\hbar}{2}\delta^{(1)}\sigma_z^{(1)} + \frac{\hbar}{2}\delta^{(2)}\sigma_z^{(2)} - \frac{\hbar}{2}\omega_1^{(1)}\sigma_x^{(1)} - \frac{\hbar}{2}\omega_1^{(2)}\sigma_x^{(2)} + \hbar\omega_{12}\sigma_z^{(1)} \otimes \sigma_z^{(2)}$$

où $\delta^{(i)}$ est le désaccord et $\omega_1^{(i)}$ la fréquence de Rabi pour le spin (i). La différence

$$\delta^{(1)} - \delta^{(2)} = \gamma(B_0^{(1)} - B_0^{(2)})$$

est le déplacement chimique (§ 5.2.3). Quels sont les quatre niveaux d'énergie en l'absence de champ de radiofréquences ($\omega_1^{(1)} = \omega_1^{(2)} = 0$) ? Introduisons l'opérateur Σ_z, composante z du spin total

$$\Sigma_z = \frac{1}{2}\,(\sigma_z^{(1)} + \sigma_z^{(1)}).$$

On peut montrer que les transitions permises correspondent à $\Delta\Sigma_z = \pm 1$, tandis que $\Delta\Sigma_z = \pm 2$ et $\Delta\Sigma_z = 0$ sont interdites. Montrer que les quatre fréquences qui apparaissent dans le signal RMN sont

$$\delta^{(1)} + \omega_{12} \qquad \delta^{(1)} - \omega_{12} \qquad \delta^{(2)} + \omega_{12} \qquad \delta^{(2)} - \omega_{12}.$$

Tracer qualitativement le schéma de niveaux et comparer avec la figure 5.10.

11.6.5 Non unicité de la préparation de l'opérateur statistique pour le spin 1/2

Soit O le centre de la sphère de Poincaré-Bloch et $\vec{b} = \vec{OM}$ le vecteur de Bloch (figure 11.1). Traçons une corde de la sphère passant par l'extrémité de $\vec{b}$. Cette corde coupe la sphère en deux points P_1 et P_2 ; définissons les vecteurs unitaires

$$\hat{n}_1 = \vec{OP}_1, \qquad \hat{n}_2 = \vec{OP}_2.$$

Montrer que le vecteur de Bloch peut s'écrire

$$\vec{b} = \hat{n}_1 + \lambda(\hat{n}_2 - \hat{n}_1) = (1-\lambda)\hat{n}_1 + \lambda\hat{n}_2, \qquad 0 < \lambda < 1$$

Montrer que la matrice statistique définie par ce vecteur $\vec{b}$ est

$$\rho = \frac{1}{2}(1-\lambda)\left(I + \vec{\sigma}\cdot\hat{n}_1\right) + \frac{1}{2}\,\lambda\left(I + \vec{\sigma}\cdot\hat{n}_2\right). \tag{11.147}$$

Montrer que l'on peut utiliser la préparation suivante : probabilité $\mathsf{p}_1 = (1-\lambda)$ pour l'état $|+,\hat{n}_1\rangle$ et probabilité $\mathsf{p}_2 = \lambda$ pour l'état $|+,\hat{n}_2\rangle$

$$\rho = \mathsf{p}_1|+,\hat{n}_1\rangle\langle+,\hat{n}_1| + \mathsf{p}_2|+,\hat{n}_2\rangle\langle+,\hat{n}_2|.$$

En déduire qu'il existe une infinité de préparations de ρ.

11.6.6 Inégalité de Wigner

Il existe plusieurs inégalités de type Bell en plus de l'inégalité BCHSH. Cet exercice examine une inégalité due à Wigner [1970], qui reprend la configuration de la figure 11.4 pour deux spins 1/2 dans l'état (11.43). Soit p_{ij}, la probabilité jointe que le spin $\mathcal{A}$ soit orienté dans la direction $\theta_i = \theta_A$ et le spin $\mathcal{B}$ dans la direction $\theta_j = \theta_B$. Nous avons montré en (11.46) que

$$\mathsf{p}_{ij} = \frac{1}{2}\sin^2(\theta_i - \theta_j) = \frac{1}{2}\sin^2\theta_{ij}$$

Nous notons σ les matrices de Pauli du spin $\mathcal{A}$ et τ celles du spin $\mathcal{B}$. Considérons les opérateurs

$$\sigma_i = \sigma(\theta_i) = \sigma_z\cos\theta_i + \sigma_x\sin\theta_i \qquad \tau_j = \tau(\theta_j) = \tau_z\cos\theta_j + \tau_x\sin\theta_j$$

Introduisons les 9 couples de valeurs pour $i, j = 1,\ 2,\ 3$

$$c_1 = \sigma_1\tau_1, c_2 = \sigma_1\tau_2, \ldots, c_9 = \sigma_3\tau_3$$

où $\sigma_i = \pm 1$ ($\tau_i = \pm 1$), est le résultat de la mesure de $\sigma(\theta_i)$ ($\tau(\theta_j)$) pour le spin $\mathcal{A}$ ($\mathcal{B}$). Chaque variable c_i peut prendre 4 valeurs : $\sigma_k\tau_l = ++, +-, -+, --$, et la distribution de probabilité $\mathsf{p}(c_1, \ldots, c_9)$ comprend donc $4^9 = 262\ 144$ termes. Montrer que la causalité relativiste implique, qu'en fait, seules $2^6 = 64$ probabilités indépendantes sont non nulles, ce que la condition $\sigma_i = -\tau_i$ réduit à 8. Notons que nous avons implicitement admis la contrafactualité : si l'on mesure, par exemple, σ_1 et τ_1, alors $\sigma_2, \sigma_3, \tau_2$ et τ_3 sont définis. La probabilité jointe p_{13} pour que le spin $\mathcal{A}$ soit orienté suivant θ_1 et le spin $\mathcal{B}$ suivant θ_3 est

$$\begin{aligned}\mathsf{p}_{13} &= \sum_{\sigma_2,\sigma_3}\sum_{\tau_1,\tau_2} \mathsf{p}(+, \sigma_2, \sigma_3; \tau_1, \tau_2, +) \\ &= \mathsf{p}(++-;--+) + \mathsf{p}(+--;-++)\end{aligned}$$

Donner l'expression analogue pour p_{12} et p_{23} et en déduire l'inégalité

$$\mathsf{p}_{12} + \mathsf{p}_{23} > \mathsf{p}_{13}$$

Trouver une configuration d'angles pour laquelle cette inégalité n'est pas satisfaite.

11.6.7 États de Hardy

Les *états de Hardy* donnent un exemple frappant de l'impossibilité de rendre compte de corrélations quantiques par une distribution de probabilité classique. Pour définir les états de Hardy, considérons deux bases incompatibles pour deux spins 1/2, (A, α) pour $\mathcal{A}$ et (B, β) pour $\mathcal{B}$, par exemple $A = \sigma_{zA}, \alpha = \sigma_{xA},\ B = \sigma_{zB}, \beta = \sigma_{xB}$, avec des vecteurs propres $|A_\pm\rangle, \ldots, |\beta_\pm\rangle$ et définissons l'état de Hardy $|\Phi_H\rangle$ par

$$|\Phi_H\rangle = |A_+B_+\rangle - |\alpha_+\beta_+\rangle\langle\alpha_+\beta_+|A_+B_+\rangle$$

Cet état n'est pas normalisé

$$\|\Phi_H\|^2 = 1 - |\langle\alpha_+\beta_+|A_+B_+\rangle|^2$$

mais ceci n'a aucune importance pour la suite. Il existe 16 résultats possibles pour les mesures ($A_\pm = \pm 1$, etc.)

$$(A_+, B_+), \ldots, (\alpha_-\beta_-)$$

Montrer que trois configurations ont une probabilité nulle. Montrer, de plus, que la probabilité conditionnelle $\mathsf{p}(A_-|B_-)$ est non nulle

$$\mathsf{p}(A_-, B_-) \neq 0 \Longrightarrow \mathsf{p}(A_-|B_-) \neq 0$$

Si l'on suppose l'existence d'une distribution de probabilité jointe

$$\mathsf{p}(A_\pm, \alpha_\pm, B_\pm, \beta_\pm)$$

montrer par inspection qu'il existe seulement 6 p différents de zéro et en déduire

$$\mathsf{p}(A_+|B_-) = 1 \quad \text{ce qui contredit} \quad \mathsf{p}(A_-|B_-) \neq 0$$

11.6.8 Photons intriqués en polarisation

On considère deux photons partant en sens inverse, l'un (1) suivant Oz et l'autre (2) suivant $-Oz$, dans un état intriqué en polarisation

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\Big(|x\rangle_1 \otimes |y\rangle_2 - |y\rangle_1 \otimes |x\rangle_2\Big) = \frac{1}{\sqrt{2}}\left(|xy\rangle - |yx\rangle\right)$$

Les états $|x\rangle$ et $|y\rangle$ sont des états de polarisation linéaire suivant Ox et Oy.

1. Soit

$$|\theta\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle$$

l'état de polarisation linéaire suivant la direction $\hat{n}_\theta$ du plan xOy et $|\theta_\perp\rangle$ l'état de polarisation orthogonale, voir (11.69). Montrer que

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\left(|\theta\,\theta_\perp\rangle - |\theta_\perp\,\theta\rangle\right)$$

L'état $|\Phi\rangle$ est donc invariant par rotation autour de Oz.

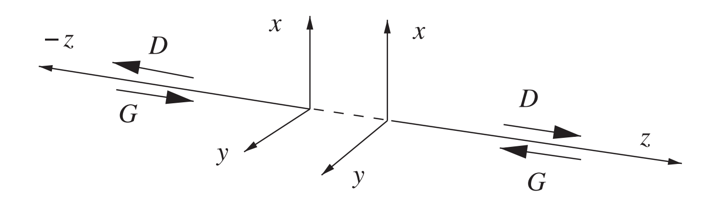


Fig. 11.20 – Configuration des polarisations des photons intriqués.

2. Écrire $|\Phi\rangle$ en fonction des états de polarisation circulaire $|D\rangle$ et $|G\rangle$ (10.25) en prenant garde à l'orientation des axes (figure 11.20) : le sens de la rotation dépend de la direction de propagation

$$|\Phi\rangle = \frac{\mathrm{i}}{\sqrt{2}}\left(|DD\rangle - |GG\rangle\right)$$

Vérifier en utilisant (11.69) que cette deuxième forme de $|\Phi\rangle$ est bien invariante par rotation autour de Oz.

3. Montrer que l'état

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|xx\rangle + |yy\rangle\right)$$

est également invariant par rotation autour de Oz. Donner son expression en fonction des états de polarisation circulaire[31].

11.6.9 Stratégies gagnantes

1. Considérons le jeu décrit dans la section 11.3.1 : Alice et Bob disposent de deux bits $a = 0, 1$ et $b = 0, 1$ et la stratégie gagnante consiste à obtenir $A \oplus B = ab$ avec le taux de succès maximum ; $A = 0, 1$ et $B = 0, 1$ sont les résultats correspondant aux entrées a et b, respectivement. Considérons une stratégie déterministe qui pour chaque choix de (a, b) donne des résultats (A, B), et telle que 3 des équations (11.68) soient satisfaites. Supposons qu'en outre Alice et Bob partagent une variable aléatoire λ pouvant prendre l'une des valeurs (1,2,3,4), cette valeur donnant le numéro de l'équation (11.68) qui n'est pas satisfaite. On définit les probabilités $\mathsf{p}(a, b; \lambda) = 1$ si la stratégie déterministe correspondant à cette valeur de λ est vérifiée et $\mathsf{p}(a, b; \lambda) = 0$ dans le cas contraire. Montrer que le taux de succès maximum du protocole probabiliste est

$$\mathrm{Max}_\lambda \frac{1}{4}\sum_{a,b} \mathsf{p}(a, b; \lambda) = \frac{3}{4}$$

2. Supposons qu'Alice et Bob partagent un état intriqué

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$$

et soit $R(\theta)$, l'opérateur de rotation autour de Oy

$$R(\theta)|0\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle \qquad R(\theta)|1\rangle = -\sin\frac{\theta}{2}|0\rangle + \cos\frac{\theta}{2}|1\rangle$$

Avant d'effectuer leur mesure, Alice et Bob appliquent cette rotation sur leurs états respectifs avec un angle θ_a pour Alice et θ_b pour Bob. Montrer que l'état après rotation est

$$|\Psi'\rangle = \frac{1}{\sqrt{2}}\Big(\cos\frac{\theta_a + \theta_b}{2}(|00\rangle + |11\rangle) + \sin\frac{\theta_a + \theta_b}{2}(|10\rangle + 01\rangle)\Big)$$

Montrer que si Alice et Bob choisissent $\theta = -\pi/8$ si $a, b = 0$ et $\theta = 3\pi/8$ pour $a, b = 1$, alors les équations (11.68) sont vérifiées avec une probabilité de $\cos^2 \pi/8$.

---

31. Les états $|\Phi\rangle$ et $|\Psi\rangle$ sont tous deux de moment angulaire nul. Si les deux photons proviennent de la désintégration d'une particule de spin 0, le choix entre les deux états dépend de la parité de la particule mère : voir l'exercice 14.5.4.

3. On passe maintenant à l'exemple à 3 particules du § 11.3.3. En utilisant un raisonnement analogue à celui de la question **1**, montrer qu'une stratégie probabiliste a un taux de succès maximum de 75 %.

4. Évaluer l'action de $H_{\mathcal{A}}\, I_{\mathcal{B}}\, H_{\mathcal{C}}$ et $H_{\mathcal{A}}\, H_{\mathcal{B}}\, I_{\mathcal{C}}$ sur l'état $|\Psi\rangle$ (11.73) et en déduire que la stratégie qui consiste à partager un état intriqué a un taux de succès maximum de 100 %.

11.6.10 États de Bell et mesure de Bell

On définit les 4 états de Bell (11.137)

$$|\Psi^{(\pm)}_{AB}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$$
$$|\Phi^{(\pm)}_{AB}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$$

1. Écrire l'expression de $|01\rangle, |10\rangle, |00\rangle, |11\rangle$ en fonction des états de Bell.

2. Utiliser ces relations pour obtenir (11.145).

3. Une mesure de Bell est schématisée sur la figure 11.15. Calculer l'action de l'opérateur (cNOT) × (H_A) sur les 4 états de Bell. Montrer que

$$(\mathrm{cNOT}) \times (\mathrm{H}_A)|\Psi^{(+)}\rangle = |0_A 1_B\rangle \qquad (\mathrm{cNOT}) \times (\mathrm{H}_A)|\Psi^{(-)}\rangle = |1_A 1_B\rangle$$

et en déduire que la mesure des qubits A et B projette sur l'un des 4 états de Bell.

11.6.11 États GHZ

On suppose qu'une particule instable se désintègre en trois particules identiques de spin 1/2, $\mathcal{A}, \mathcal{B}$ et $\mathcal{C}$ émises dans un plan, dans une configuration où les trois impulsions font entre elles un angle de $2\pi/3$ et dans l'état intriqué de spin

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(|+++\rangle - |---\rangle\Big)$$

Trois expérimentateurs, Alice (a), Bob (b) et Carla (c) peuvent mesurer la composante du spin suivant une direction perpendiculaire à la direction de propagation de chaque particule (figure 11.21). Le plan des impulsions est le plan horizontal, l'axe Oz est choisi le long de la direction de propagation (il dépend donc de la particule), l'axe Oy est vertical et $\hat{x} = \hat{y} \times \hat{z}$.

1. On considère les trois opérateurs

$$\Sigma_a = \sigma_{ax}\sigma_{by}\sigma_{cy} \quad \Sigma_b = \sigma_{ay}\sigma_{bx}\sigma_{cy} \quad \Sigma_c = \sigma_{ay}\sigma_{by}\sigma_{cx}$$

Attention : mesurer Σ_a par exemple n'est pas équivalent à mesurer séparément σ_{ax}, σ_{by} et σ_{cy} et faire le produit des résultats ! En utilisant l'action des opérateurs σ_x et σ_y sur les états $|+\rangle$ et $|-\rangle$, montrer que

$$\Sigma_a|\Psi\rangle = \Sigma_b|\Psi\rangle = \Sigma_c|\Psi\rangle = |\Psi\rangle$$

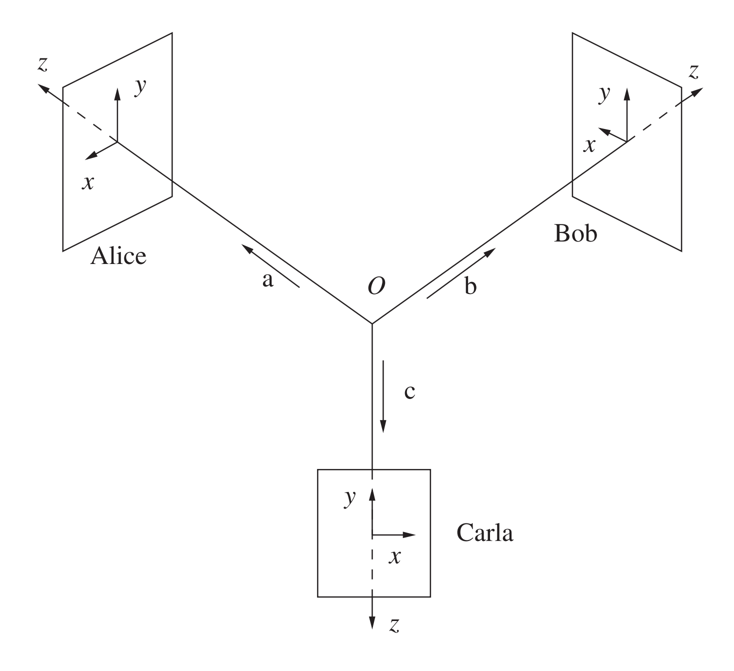


FIG. 11.21 – Configuration d'une expérience de type GHZ.

Soit $A_x = \pm 1$, le résultat de la mesure de σ_x pour le spin $\mathcal{A}$ par Alice, ..., C_y le résultat de la mesure de σ_y pour le spin $\mathcal{C}$ par Carla. Montrer que

$$A_x B_y C_y = +1 \qquad A_y B_x C_y = +1 \qquad A_y B_y C_x = +1$$

2. Supposons qu'au lieu de décider de mesurer une composante suivant Ox et deux composantes suivant Oy, les trois expérimentateurs décident tous de mesurer σ_x. Cette décision peut être prise alors que les trois spins sont en vol et ne peuvent plus communiquer entre eux. Montrer que

$$\sigma_{ax}\sigma_{bx}\sigma_{cx}|\Psi\rangle \equiv \Sigma|\Psi\rangle = -|\Psi\rangle$$

et par conséquent $A_x B_x C_x = -1$. Monter que ce résultat est incompatible avec celui de la question **1**.

Montrer que $\Sigma = -\Sigma_a\Sigma_b\Sigma_c$ et en déduire que Σ, Σ_a, Σ_b et Σ_c peuvent être mesurés simultanément. Discuter la façon dont la contextualité se manifeste dans cette expérience.

11.6.12 Théorème de non-clonage quantique

Effectuer la trace partielle pour montrer (11.113) et en déduire l'expression (11.118) des fidélités.

11.6.13 Discrimination entre deux états non orthogonaux

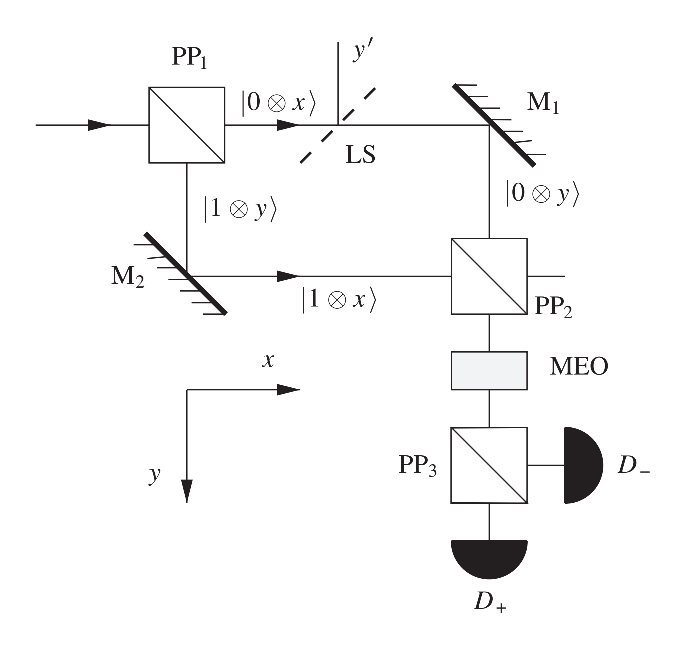


Fig. 11.22 – Expérience distinguant deux états non orthogonaux. Le photon incident rencontre un premier prisme polarisant PP_1 qui fabrique deux états $|0 \otimes x\rangle$ et $|1 \otimes y\rangle$. La lame semi-transparente LS a un coefficient de transmission t. L'action combinée du modulateur électrooptique MEO et du prisme polarisant PP_3 sépare les états $|+\rangle$ et $|-\rangle$. D'après Scarani [2011].

On prépare avec une probabilité 1/2 une suite d'états $|\psi_\pm\rangle$ d'un espace à deux dimensions $\mathcal{H}_A$ (par exemple un spin 1/2 ou la polarisation d'un photon)

$$|\psi_\pm\rangle = \cos\theta\,|0_A\rangle \pm \sin\theta\,|1_A\rangle$$

où $\{|0_A\rangle, |1_A\rangle\}$ est une base orthonormée de cet espace et $-\pi/2 \le \theta \le \pi/2$; le produit scalaire $\langle\psi_-|\psi_+\rangle = \cos 2\theta$. On introduit un système auxiliaire B également à deux niveaux dans l'état $|0_B\rangle$ et on forme le produit tensoriel

$$|\psi_\pm \otimes 0_B\rangle = \cos\theta\,|0_A \otimes 0_B\rangle \pm \sin\theta\,|1_A \otimes 0_B\rangle$$

On applique ensuite une transformation unitaire dans l'espace $\{|0_A \otimes 0_B\rangle, |1_A \otimes 1_B\rangle\}$

$$\cos\theta\,|0_A \otimes 0_B\rangle \to \sin\theta\,|0_A \otimes 0_B\rangle + \cos^{1/2} 2\theta|1_A \otimes 1_B\rangle$$

1. Montrer que cette transformation est bien unitaire et qu'après cette transformation, $|\psi_\pm \otimes 0_B\rangle$ devient

$$|\Phi_\pm\rangle = \sqrt{2}\,\sin\theta\,|\pm_A \otimes 0_B\rangle + \cos^{1/2} 2\theta|1_A \otimes 1_B\rangle$$

avec

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$$

2. En remarquant que l'état $|0_B\rangle$ est corrélé à $|\pm_A\rangle$, en déduire la probabilité de pouvoir discriminer entre les deux états $|\psi_\pm\rangle$ et relier cette probabilité à $\langle\psi_-|\psi_+\rangle$.

3. On réalise le dispositif de la figure 11.22, avec des photons qui rencontrent des prismes polarisants PP et une lame semi-transparente LS dont le coefficient de transmission est t (Scarani [2011]) ; $|0\rangle$ et $|1\rangle$ représentent des états de polarisation horizontaux et verticaux des photons, qui correspondent au système A de la question **1**. Les prismes PP_1 et PP_2 transmettent la polarisation $|0\rangle$ et réfléchissent la polarisation $|1\rangle$, tandis que $|x\rangle$, $|y\rangle$ et $|y'\rangle$ désignent des états de propagation suivant les directions Ox, Oy et y' correspondant au sytème B de la question **1**. Montrer que l'évolution de $|\psi_\pm\rangle$ intriqué avec un état de propagation $|x\rangle$ est donné par

$$\begin{aligned}
|\psi_\pm \otimes x\rangle \quad &\overset{PP_1}{\rightarrow} \quad \cos\theta|0\otimes x\rangle \pm \sin\theta|1\otimes y\rangle \\
&\overset{LS}{\rightarrow} \quad \sqrt{t}\,\cos\theta\,|0\otimes y\rangle \pm \sin\theta\,|1\otimes y\rangle + \mathrm{i}\sqrt{1-t}\,\cos\theta\,|0\otimes y'\rangle \\
&\overset{\text{miroirs}}{\rightarrow} \quad (\sqrt{t}\,\cos\theta\,|0\otimes y\rangle \pm \sin\theta\,|1\otimes x\rangle) + \mathrm{i}\sqrt{1-t}\,\cos\theta\,|0\otimes y'\rangle \\
&\overset{PP2}{\rightarrow} \quad \left(\sqrt{t}\,\cos\theta\,|0\rangle \pm \sin\theta\,|1\rangle\right) \otimes |y\rangle + \mathrm{i}\sqrt{1-t}\,\cos\theta\,|0\otimes y'\rangle
\end{aligned}$$

Quelle valeur de t faut-il choisir pour que $|y\rangle$ soit en facteur d'un produit tensoriel avec l'un de deux états de polarisation orthogonaux $|\pm\rangle$? Montrer que l'action combinée du modulateur électro-optique MEO et du prisme polarisant PP_3 permet de distinguer les états $|+\rangle$ et $|-\rangle$.

11.6.14 Interférences des temps d'émission

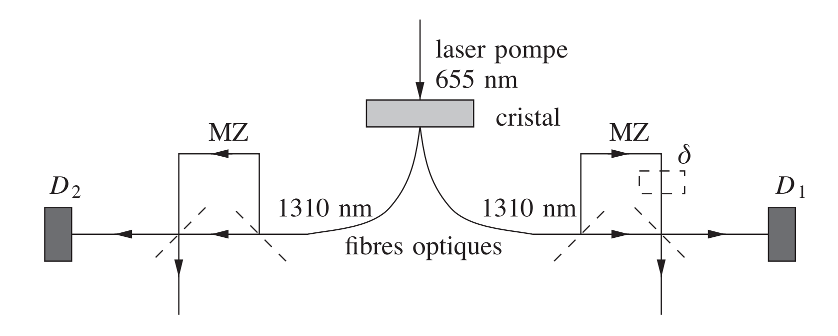


FIG. 11.23 – Interférences des temps d'émission.

Dans une expérience réalisée par une collaboration Nice-Genève (Tanzilli *et al.* [2002]), un faisceau laser (laser de pompe) incident de longueur d'onde

$\lambda = 655\,\text{nm}$ arrive sur un cristal non linéaire (figure 11.23). Une fraction des photons incidents est convertie en paires de photons de longueur d'onde $2\lambda = 1310\,\text{nm}$, chaque photon partant dans une des deux fibres optiques et traversant ensuite un interféromètre de Mach-Zehnder (MZ) (*cf.* § 1.4.5). Ces interféromètres ont un bras court et un bras long, la différence entre les deux bras étant $\Delta l = 20\,\text{cm}$. Une lame permet de faire varier le chemin optique de δ sur le bras long de l'interféromètre de droite. La longueur de cohérence (§ 5.4.2) $l_{\text{coh}} \simeq 40\,\mu\text{m}$ des photons convertis est très petite par rapport à Δl : $l_{\text{coh}} \ll \Delta l$ (alors que la longueur de cohérence du laser de pompe est voisine de 100 m).

1. On fait varier la phase δ sur le bras long de l'interféromètre de droite. Montrer que le taux de comptage des photons par le détecteur D_1 est indépendant de δ.

2. On détecte les deux photons en coïncidence dans D_1 et D_2, avec une fenêtre de coïncidence de l'ordre de 0.1 ns ; comme le faisceau pompe est continu, on ne dispose d'aucune information sur le temps de génération d'une paire de photons. Montrer qu'il n'est pas possible de distinguer entre les deux chemins court-court et long-long suivis par les photons. En déduire que si l'on fait varier δ, on obtient une variation sinusoïdale du taux de comptage en coïncidence, mais que les taux de détection individuels dans D_1 et D_2 restent indépendants de δ. Suggestion : montrer que si l'on supprime les deux diviseurs de faisceau du MZ de gauche, on peut déduire une information sur le trajet suivi par le photon de droite. Que se passe-t-il si l'on supprime l'ensemble du dispositif de gauche (MZ et détecteurs) ?

11.6.15 Calcul quantique avec des ions piégés

1. Les ions piégés sont peut-être une solution d'avenir pour la construction d'un ordinateur quantique. Dans une expérience réalisée par un groupe d'Innsbruck (Schmid-Kaler *et al.* [2003], Blatt [2004]), des ions $^{40}\text{Ca}^+$ sont confinés dans un piège approximativement unidimensionnel et harmonique. L'état fondamental $S_{1/2} = |g\rangle$ est identifié à l'état $|0\rangle$ du calcul quantique (11.5.2), et l'état excité $D_{5/2} = |e\rangle$ avec $|1\rangle$. L'état excité est métastable, de vie moyenne très longue ($\sim 1\,\text{s}$) parce que la transition $D_{5/2} \to S_{1/2}$ est quadrupolaire électrique. Considérons d'abord un ion unique dans le piège. Son hamiltonien approché est

$$H_{\text{trap}} = \frac{1}{2M}\,p_z^2 + \frac{1}{2}M\omega_z^2 z^2$$

où M est la masse de l'ion et ω_z la fréquence du piège. En absence d'un champ appliqué, le hamiltonien est la somme du hamiltonien des états internes de l'ion et de l'oscillateur harmonique du piège

$$H_0 = -\frac{1}{2}\,\hbar\omega_0\sigma_z + \hbar\omega_z a^\dagger a$$

où ω_0 est la fréquence de la transition $|0\rangle \leftrightarrow |1\rangle$. On applique à l'ion le champ d'un laser lorsque $t > 0$

$$\vec{E} = E_1 \hat{x} \cos(\omega t - kz - \phi)$$

La fréquence de Rabi est notée ω_1. Le couplage entre le champ et l'ion est

$$H_{\rm int} = -\frac{1}{2}\, \hbar \sigma_x \cos(\omega t - kz - \phi)$$

Ce hamiltonien s'écrit dans le point de vue de l'interaction (§ 4.2.5)

$$\tilde{H}_{\rm int}(t) = \mathrm{e}^{\,\mathrm{i}H_0 t}\, H_{\rm int}(t)\, \mathrm{e}^{-\mathrm{i}H_0 t}$$

Montrer que l'approximation séculaire conduit au hamiltonien

$$\tilde{H}_{\rm int} \simeq -\frac{1}{2}\, \omega_1 \left[\sigma_+\, \mathrm{e}^{\,\mathrm{i}(\delta t - \phi)}\, \mathrm{e}^{-\mathrm{i}k\tilde{z}} + \sigma_-\, \mathrm{e}^{-\mathrm{i}(\delta t - \phi)}\, \mathrm{e}^{\,\mathrm{i}k\tilde{z}}\right]$$

où $\delta = \omega - \omega_0$ est le désaccord. Si $\delta = 0$, ce hamiltonien est indépendant du temps

$$\tilde{H}_1 \simeq -\frac{1}{2}\, \omega_1 \left[\sigma_+\, \mathrm{e}^{-\mathrm{i}\phi}\, \mathrm{e}^{-\mathrm{i}k\tilde{z}} + \sigma_-\, \mathrm{e}^{\,\mathrm{i}\phi}\, \mathrm{e}^{\,\mathrm{i}k\tilde{z}}\right]$$

Comme

$$\tilde{z} = \sqrt{\frac{\hbar}{2M\omega_z}}\, \left(a + a^\dagger\right)$$

$\exp(\pm \mathrm{i}k\tilde{z})$ couple les niveaux internes $|0\rangle$ and $|1\rangle$ aux niveaux de vibration dans le piège. Les niveaux internes seront étiquetés n, $n = 0,\ 1$, les niveaux de vibration m, $m = 0,\ 1,\ 2,\ \ldots$, et l'état produit $|n, m\rangle$

2. Définissons le paramètre sans dimension de Lamb-Dicke η par

$$\eta = k\sqrt{\frac{\hbar}{2M\omega_z}}$$

Quelle est l'interpretation physique de η ? Soit m et $m + m'$, deux niveaux de l'oscillateur harmonique. Montrer que la fréquence de Rabi $\omega_1^{m \to m+m'}$ est donnée par

$$\omega_1^{m \to m+m'} = \omega_1 |\langle m + m'|\mathrm{e}^{\,\mathrm{i}\eta(a+a^*)}|m\rangle|$$

où η est le paramètre de Lamb-Dicke. Les transitions correspondant aux fréquences $\omega = \omega_0 + \omega_z$ (resp. $\omega = \omega_0 - \omega_z$) sont appelées bande latérale bleue (resp. bande latérale rouge), tandis que les transitions avec $\omega = \omega_0$ sont appelées transitions porteuses. Nous supposons aussi que $\eta \ll 1$ et nous nous plaçons au premier ordre en η. Montrer qu'à cette approximation et pour $m' = \pm 1$, on a pour les bandes latérales bleues et rouges

$$\begin{aligned} \omega_1^{m \to m+1} &\simeq \eta\sqrt{m+1}\,\omega_1 = \omega_1^+ \quad \text{(bleue)} \\ \omega_1^{m \to m-1} &\simeq \eta\sqrt{m}\,\omega_1 = \omega_1^- \quad \text{(rouge)} \end{aligned}$$

Écrire l'expression de $\tilde{H}_{\rm int}$ pour les deux bandes latérales, et montrer que pour la bande latérale bleue

$$\tilde{H}^{+}_{\rm int} = \frac{\rm i}{2}\,\eta\,\omega_1\sqrt{m+1}\Big[\sigma_+ a_b\,{\rm e}^{-{\rm i}\phi} - \sigma_- a_b^\dagger {\rm e}^{\,{\rm i}\phi}\Big]$$

tandis que pour la bande latérale rouge

$$\tilde{H}^{-}_{\rm int} = \frac{\rm i}{2}\,\eta\,\omega_1\sqrt{m}\,\Big[\sigma_+ a_r^\dagger {\rm e}^{-{\rm i}\phi} - \sigma_- a_r\,{\rm e}^{\,{\rm i}\phi}\Big]$$

Les opérateurs $a_b \ldots a_r^\dagger$ sont définis de façon à conserver la norme des vecteurs d'état

$$a_b = \frac{a}{\sqrt{m+1}} \qquad a_b^\dagger = \frac{a^\dagger}{\sqrt{m+1}} \qquad a_r = \frac{a}{\sqrt{m}} \qquad a_r^\dagger = \frac{a^\dagger}{\sqrt{m}}$$

On se limite au cas $m = 1$. Quels sont les opérateurs de rotation sur les deux bandes $R^{\pm}(\theta, \phi)$, où $\theta = -\omega_1^{\pm} t$?

3. En plus des niveaux $|0,0\rangle$, $|0,1\rangle$, $|1,0\rangle$ et $|1,1\rangle$, on utilise également le niveau $|1,2\rangle$. Dessiner le schéma des niveaux et identifier les transitions de bande latérale bleue $|0,0\rangle \leftrightarrow |1,1\rangle$ et $|0,1\rangle \leftrightarrow |1,2\rangle$. Montrer que l'opérateur de rotation $R^{+}_{\alpha\beta}$ défini par

$$R^{+}_{\alpha\beta} = R^{+}(\alpha, \pi/2)\, R^{+}(\beta, 0)\, R^{+}(\alpha, \pi/2)\, R^{+}(\beta, 0)$$

est égal à $-I$ pour $\alpha = \pi$, β quelconque, ou $\beta = \pi$, α quelconque. $R^{\pm}(\theta, \phi)$ est une rotation d'angle θ autour d'un axe situé dans le plan xOy, faisant un angle ϕ avec l'axe Ox et utilisant la bande latérale bleue $(+)$ ou rouge $(-)$. Compte tenu de ce que la fréquence de Rabi pour la transition $|0,1\rangle \leftrightarrow |1,2\rangle$ vaut $\sqrt{2}$ fois celle pour la transition $|0,0\rangle \leftrightarrow |1,1\rangle$, comment peut-on choisir α et β de telle sorte que $R^{+}_{\alpha\beta} = -I$ pour les deux transitions ? En déduire la séquence des 4 impulsions et leur durée de telle sorte que le résultat net soit

$$|00\rangle \leftrightarrow -|0,0\rangle \qquad |0,1\rangle \leftrightarrow -|0,1\rangle \qquad |1,0\rangle \leftrightarrow +|1,0\rangle \qquad |1,1\rangle \leftrightarrow -|1,1\rangle$$

On a donc fabriqué une porte cZ (au signe près)[32].

4. Il faut maintenant “transférer” la porte cZ vers la base de calcul des états $|n_1, n_2\rangle$, $n_1, n_2 = 0, 1$ étant les états fondamentaux et excités des deux ions. Montrer que l'on obtient le résultat souhaité en prenant en sandwich l'opérateur de rotation $R^{+(1)}_{\alpha\beta}$ sur l'ion numéro 1 utilisant la bande latérale bleue entre deux rotations de π sur l'ion numéro 2 utilisant la bande latérale rouge

$$\big[R^{-(2)}(\pi, \pi/2)\big]\, R^{+(1)}_{\alpha\beta}\, \big[R^{-(2)}(-\pi, \pi/2)\big]$$

Une opération un peu plus complexe permet de construire une porte cNOT.

32. La porte cZ se déduit de la porte cNOT (11.125) par $\sigma_x \to \sigma_z$. En fait, cNOT $=$ H$^{\otimes 2}$cZ H$^{\otimes 2}$, où H est la porte de Hadamard (11.122).

11.7 Bibliographie

L'opérateur statistique est traité dans Messiah [1959], chapitres VII et VIII ou Cohen-Tannoudji *et al.* [1973], compléments E_{III} et E_{IV}. Deux références plus récentes sont Isham [1995], chapitre 6 ou Basdevant et Dalibard [2001], annexe D. Une discussion approfondie des concepts fondamentaux est donnée par Mermin [1998]. Pour les applications de l'opérateur statistique à la mécanique statistique et les propriétés de l'entropie de von Neumann, on pourra consulter Balian [1991], chapitres 2 à 5, Diu *et al.* [1990], chapitre 2 ou Le Bellac *et al.* [2004], chapitre 2. Les applications de l'opérateur statistique à la RMN sont discutées par exemple par Levitt [2001], chapitre 10. Pour une revue générale des sujets abordés dans ce chapitre et au delà, voir Scarani [2011] et Laloë [2011] ; un exposé élémentaire se trouve dans Zeilinger [2010] ou Gisin [2012], qui traite la non-localité très en détail. Il existe de nombreux exposés sur les inégalités de Bell, parmi lesquels on peut recommander ceux de Bell [2004], Peres [1993], chapitres 6 et 7, Isham [1995], chapitres 8 et 9, Mermin [1993] et Norsen [2006]. On trouvera dans ces références une discussion de la contextualité et des théorèmes de Gleason et de Kochen-Specker. Les cours de Cohen-Tannoudji [1989] et [1990] contiennent un exposé très complet de la théorie de la mesure et de la décohérence ; voir aussi Zurek [1991] et [2003], Leggett [2002a], [2002b], [2005] et [2008], Joos et Zeh [1985], Schlossauer [2004] et d'Espagnat [1995]. L'exposé du § 6.4.5 suit Ballentine [1990] ; voir également Cohen-Tannoudji [1990]. Pour une excellente introduction au calcul quantique, on pourra consulter le livre de Nielsen et Chuang [2001] ; des livres plus récents (et plus courts !) sont ceux de Stolze et Suter [2004], Le Bellac [2006], Mermin [2010] et Rieffel et Polak [2011]. Le théorème de non-clonage quantique et ses applications sont traités en détail par Scarani *et al.* [2005] ; voir également Buzek et Hillery [1996]. On trouvera une version grand public de la téléportation dans Zeilinger [2000]. Les articles "historiques" (antérieurs à 1982) ont été rassemblés dans un ouvrage édité par Wheeler et Zurek [1983]. On y trouvera en particulier les article originaux de EPR (Einstein *et al.* [1935]) et de Bohr [1935]. Les corrélations EPR ont été utilisées dans l'expérience montrant la non-invariance par renversement du temps dans la physique des mésons-B : Lees *et al.* [2012], Schwarzschild [2012].

Annexes

A Théorème de Wigner et renversement du temps

Dans cette annexe, nous allons donner la démonstration du théorème de Wigner énoncé au § 7.1.2 et examiner l'invariance par rapport au renversement du sens du temps (ou simplement renversement du temps), qui est particulière car l'opérateur qui réalise la symétrie dans l'espace de Hilbert $\mathcal{H}$ est antiunitaire, et non unitaire comme dans tous les cas que nous avons rencontrés jusqu'à présent. Rappelons la définition du § 7.1.1 d'un rayon de l'espace de Hilbert : *un rayon est un vecteur à un facteur de phase près.* Deux vecteurs unitaires φ et $\overline{\varphi}$ qui diffèrent par un facteur de phase : $\varphi = \exp(\mathrm{i}\alpha)\,\overline{\varphi}$ appartiennent à une même classe d'équivalence, qui est précisément un rayon $\tilde{\varphi}$ de $\mathcal{H}$. Comme le module du produit scalaire est indépendant du représentant dans la classe d'équivalence

$$|(\overline{\varphi}, \overline{\chi})| = |(\varphi, \chi)|$$

le *module* du produit scalaire de deux rayons $\tilde{\varphi}$ et $\tilde{\chi}$ est bien défini en choisissant deux représentants arbitraires dans chaque classe d'équivalence

$$|(\tilde{\varphi}, \tilde{\chi})| = |(\varphi, \chi)| \tag{A.1}$$

mais bien évidemment cela n'a pas de sens de parler du produit scalaire de deux rayons. Nous utilisons la notation $(\bullet, \bullet)$ pour le produit scalaire, afin d'éviter les ambiguïtés de la notation de Dirac, qui seraient particulièrement gênantes dans cette annexe.

Soit dans $\mathcal{H}$ une correspondance entre *rayons*

$$\tilde{\varphi} \to T\tilde{\varphi} \tag{A.2}$$

telle que le module du produit scalaire soit invariant

$$|(\tilde{\varphi}, \tilde{\chi})| = |(T\tilde{\varphi}, T\tilde{\chi})| \tag{A.3}$$

Le théorème de Wigner énonce qu'il est toujours possible de choisir les phases des vecteurs de telle sorte que la correspondance entre *rayons* devienne une correspondance entre *vecteurs*

$$\varphi \to U\varphi \qquad |(U\varphi, U\chi)| = |(\varphi, \chi)| \tag{A.4}$$

où la transformation U est soit linéaire unitaire

$$(U\varphi, U\chi) = (\varphi, \chi) \tag{A.5}$$

soit antilinéaire unitaire (= antiunitaire)

$$(U\varphi, U\chi) = (\chi, \varphi) = (\varphi, \chi)^* \tag{A.6}$$

A.1 Démonstration du théorème

Soit $\{\chi_i\}, i = 1, \dots N$ une base orthonormée de $\mathcal{H}$ supposé de dimension N, $(\chi_i, \chi_k) = \delta_{ik}$. Nous allons faire jouer un rôle particulier au premier vecteur de base : *par convention, les indices i et k varieront entre 1 et N et les indices j et l entre 2 et N.* Choisissons un représentant $\chi_1'' \equiv \chi_1'$ dans la classe de $T\tilde{\chi}_1$ et un représentant χ_j'' dans la classe de $T\tilde{\chi}_j$, $j = 2, \dots, N$. D'après (A.3), l'ensemble $\{\chi_1'', \chi_j''\}$ forme aussi une base de $\mathcal{H}$ car

$$|(\chi_i'', \chi_k'')| = |(\chi_i, \chi_k)| = \delta_{ik}$$

Considérons l'ensemble des vecteurs φ_j

$$\varphi_j = \chi_1 + \chi_j \qquad j = 2, \dots, N \tag{A.7}$$

et soit $T\tilde{\varphi}_j$, le transformé du rayon $\tilde{\varphi}_j$. Si φ_j'' est un représentant de $T\tilde{\varphi}_j$, nous aurons

$$|(\chi_1', \varphi_j'')| = |(\chi_1, \varphi_j)| = 1$$
$$|(\chi_j'', \varphi_l'')| = |(\chi_j, \varphi_l)| = \delta_{jl}$$

Un représentant φ_j'' de $T\tilde{\varphi}_j$ aura donc des composantes uniquement suivant χ_1' et χ_j''

$$\varphi_j'' = c_j \chi_1' + d_j \chi_j''$$

et ces composantes seront de module unité : $|c_j| = |d_j| = 1$. On peut maintenant choisir des représentants φ_j' et χ_j'

$$\varphi_j' = \frac{1}{c_j}\, \varphi_j'' \qquad \chi_j' = \frac{d_j}{c_j}\, \chi_j'' \tag{A.8}$$

de sorte que

$$\varphi_j' = \frac{1}{c_j}(c_j \chi_1' + d_j \chi_j'') = \chi_1' + \chi_j' \tag{A.9}$$

Nous avons donc défini une application sur des *vecteurs* de $\mathcal{H}$

$$\chi_1 + \chi_j \to (\chi_1 + \chi_j)' = \chi_1' + \chi_j'$$

telle que $\chi_1' \in T\tilde{\chi}_1$, $\chi_j' \in T\tilde{\chi}_j$ et $\chi_1' + \chi_j' \in T(\widetilde{\chi_1 + \chi_j})$. Essayons maintenant de déterminer s'il est possible qu'un vecteur arbitraire ψ se transforme suivant

$$\psi = \sum_{k=1}^{N} c_k \chi_k \to \psi' = \sum_{k=1}^{N} c_k' \chi_k'$$

Si une telle loi de transformation est valide, nous devons avoir, d'une part,

$$|c_k'| = |(\chi_k', \psi')| = |(\chi_k, \psi)| = |c_k|$$

et, d'autre part,

$$(\chi_1 + \chi_j, \psi) = c_1 + c_j \qquad (\chi_1' + \chi_j', \psi') = c_1' + c_j'$$

ce qui implique, d'après (A.3), que

$$|c_1 + c_j| = |c_1' + c_j'| \tag{A.10}$$

Les deux couples de nombres complexes (c_1, c_j) et (c_1', c_j') doivent être tels que $|c_1| = |c_1'|$ et $|c_j| = |c_j'|$ et de plus vérifier (A.10). Posons

$$c_1 = |c_1|\,\mathrm{e}^{\mathrm{i}\theta_1} \qquad c_j = |c_j|\,\mathrm{e}^{\mathrm{i}\theta_j}$$
$$c_1' = |c_1'|\,\mathrm{e}^{\mathrm{i}\theta_1'} \qquad c_j' = |c_j'|\,\mathrm{e}^{\mathrm{i}\theta_j'}$$

Les angles (θ_1, θ_j) et (θ_1', θ_j') sont liés par l'équation

$$\cos(\theta_1 - \theta_j) = \cos(\theta_1' - \theta_j') \tag{A.11}$$

qui a deux solutions

$$\theta_1 - \theta_j = \theta_1' - \theta_j' \tag{A.12}$$
$$\theta_1 - \theta_j = -(\theta_1' - \theta_j') \tag{A.13}$$

Examinons le premier cas. On peut redéfinir la phase de ψ' de telle sorte que $c_1' = c_1$ et donc $\theta_1' = \theta_1$. Dans ce cas, $\theta_j' = \theta_j$ et $c_j' = c_j$

$$\psi' = \sum_k c_k \chi_k'$$

Si l'on considère un autre vecteur $\eta = \sum_k d_k \chi_k$ avec à nouveau $d_1' = d_1$, on aura

$$(\lambda\psi + \mu\eta)' = \sum_k (\lambda c_k + \mu d_k)\chi_k' = \lambda\psi' + \mu\eta'$$

Par un choix convenable des phases, la transformation T peut être choisie linéaire, et comme elle conserve le module du produit scalaire, elle est aussi unitaire : $T \to U$ avec $U^\dagger U = UU^\dagger = I$.

Dans le second cas, on redéfinit la phase de ψ' de telle sorte que $c'_1 = c^*_1$. On a alors $c'_j = c^*_j$ et

$$\psi' = \sum_k c^*_k \chi'_k$$

Le transformé de $\lambda\psi + \mu\eta$ est alors

$$(\lambda\psi + \mu\eta)' = \left[\sum_k (\lambda c_k + \mu d_k)\chi_k\right]' = \lambda^*\psi' + \mu^*\eta' \tag{A.14}$$

et la loi de transformation du produit scalaire est

$$(\psi', \eta') = (\psi, \eta)^* = (\eta, \psi) \tag{A.15}$$

La transformation $T \to V$, où V est dite *antiunitaire* : elle est antilinéaire et conserve la norme.

La démonstration précédente est en fait, incomplète : en effet, il faudrait montrer que l'on ne peut pas avoir (A.12) pour c_j et (A.13) pour $c_l, l \neq j$. La vérification que ceci ne peut pas se produire est fastidieuse et laissée au lecteur[1] : il faut examiner le comportement du transformé d'un vecteur $\psi = \chi_1 + \chi_j + \chi_l$.

A.2 Renversement du sens du temps

En mécanique classique, l'équation de Newton

$$m\frac{\mathrm{d}^2\vec{r}(t)}{\mathrm{d}t^2} = \vec{F}(\vec{r}(t))$$

est invariante par renversement du sens du temps $t \to -t$. Posons en effet $\vec{r}\,'(t) = \vec{r}(-t)$

$$m\frac{\mathrm{d}^2\vec{r}\,'(t)}{\mathrm{d}t^2} = m\frac{\mathrm{d}^2\vec{r}(-t)}{\mathrm{d}t^2} = \vec{F}(\vec{r}(-t)) = \vec{F}(\vec{r}\,'(t))$$

On constate que $\vec{r}\,'(t)$ obéit bien aux équations de Newton. La raison en est évidemment que ces équations ne dépendent que de la dérivée seconde par rapport au temps de $\vec{r}$ et pas de la dérivée première[2]. Une

1. Voir Weinberg [1995], chapitre 2, qui détaille toutes les subtilités de la preuve.
2. Une équation du type oscillateur harmonique amorti

$$m\ddot{x} + \gamma\dot{x} + m\omega^2 x = 0$$

n'est pas invariante par renversement du sens du temps, mais la force de viscosité $-\gamma\dot{x}$ est une force effective, représentant phénoménologiquement l'effet des collisions des molécules du fluide sur la particule de masse m.

image intuitive du renversement du temps est la suivante : imaginons que nous suivions la trajectoire d'une particule de $t = -\infty$ à $t = 0$ et qu'à $t = 0$, nous renversions brutalement le sens de l'impulsion (ou de la vitesse) : $\vec{p}(0) \to -\vec{p}(0)$. Dans ces conditions, la particule va "remonter sa trajectoire", elle repassera au temps t par la position qu'elle avait au temps $-t$ avec une impulsion opposée (figure A.1)

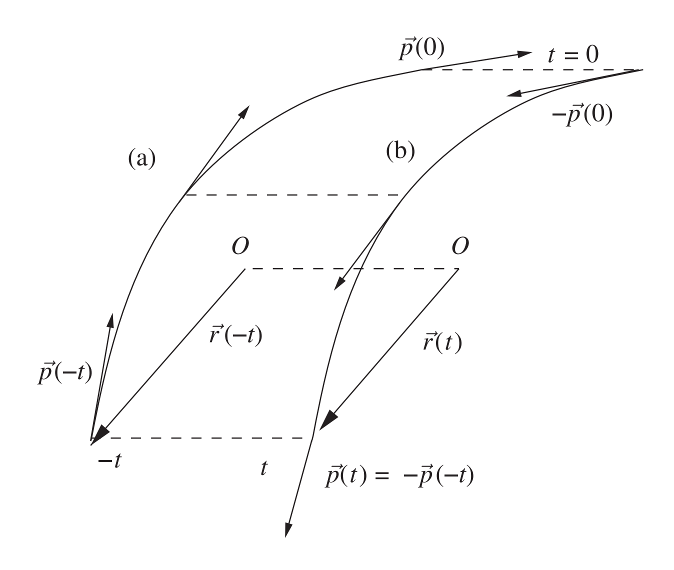


FIG. A.1 – Renversement du temps sur une trajectoire classique.

$$\vec{r}\,'(t) = \vec{r}(-t) \qquad \vec{p}\,'(-t) = -\vec{p}(t) \tag{A.16}$$

Le vecteur position $\vec{r}$ est pair par renversement du temps, et $\vec{p}$ est impair dans cette même opération. L'invariance par renversement du temps est appelée *microréversibilité*. Si l'on filme le mouvement de particules et que l'on projette le film à l'envers, la microréversibilité implique que la projection apparaît physiquement possible[3]. On sait que tel n'est pas le cas dans la vie courante, qui est fondamentalement irréversible, et il n'est pas évident[4] de comprendre comment une dynamique réversible à l'échelle microscopique peut conduire à des phénomènes irréversibles à l'échelle macroscopique.

3. On notera l'analogie avec la conservation de la parité : l'image d'une expérience dans un miroir apparaît physiquement possible si la parité est conservée.

4. Comme l'ont montré les discussions acharnées de Boltzmann avec ses contradicteurs ! Voir par exemple Balian [1991], chapitre 15 ou Le Bellac *et al.* [2004], chapitre 2.

Revenons à la mécanique quantique, en appelant Θ l'opérateur qui réalise le renversement du temps dans $\mathcal{H}$. Cet opérateur doit transformer $\vec{R}$, $\vec{P}$ et $\vec{J}$ suivant

$$\begin{aligned} \Theta\,\vec{R}\,\Theta^{-1} &= \vec{R} \\ \Theta\,\vec{P}\,\Theta^{-1} &= -\vec{P} \\ \Theta\,\vec{J}\,\Theta^{-1} &= -\vec{J} \end{aligned} \tag{A.17}$$

En effet, $\vec{J}$ doit se transformer comme $\vec{R}\times\vec{P}$, qui est impair par renversement du temps : le moment angulaire définit un sens de rotation qui est inversé par renversement du temps. L'examen de la transformation par Θ des relations de commutation canoniques montre que Θ doit être antiunitaire. Calculons de deux façons différentes un élément de matrice du commutateur $[X_i, P_j] = \mathrm{i}\,\hbar\,\delta_{ij} I$

$$\begin{aligned} (\Theta\varphi, \Theta[X_i, P_j]\psi) &= (\Theta\varphi, \Theta\mathrm{i}\,\hbar\,\delta_{ij} I\psi) = \delta_{ij}(\varphi, \mathrm{i}\,\hbar\psi)^* = -\mathrm{i}\,\hbar\,\delta_{ij}(\varphi,\psi)^* \\ &= (\Theta\varphi, \Theta[X_i, P_j]\Theta^{-1}\Theta\psi) \\ &= (\Theta\varphi, -\mathrm{i}\hbar\delta_{ij} I\Theta\psi) = -\mathrm{i}\hbar\delta_{ij}(\varphi,\psi)^* \end{aligned}$$

où nous avons utilisé dans la seconde ligne les lois de transformation (A.17) de X_i et P_j

$$\Theta[X_i, P_j]\Theta^{-1} = -[X_i, P_j]$$

Les deux lignes de l'équation précédente sont compatibles, ce qui ne serait pas le cas si la transformation Θ était unitaire.

Il existe un autre argument très instructif prouvant le caractère antiunitaire de Θ. Soit $\varphi(t)$, le vecteur d'état d'un système quantique au temps t, $\varphi = \varphi(t=0)$ son état au temps $t = 0$

$$\varphi(t) = \exp\left(-\frac{\mathrm{i}}{\hbar}\,Ht\right)\varphi$$

L'invariance par rapport au renversement du temps implique que l'état transformé de $\varphi(-t)$ par renversement du temps, $\Theta\varphi(-t)$, coïncide avec l'état obtenu par évolution temporelle de $\Theta\varphi(t=0)$

$$\Theta\varphi(-t) = \exp\left(-\frac{\mathrm{i}}{\hbar}\,Ht\right)\Theta\varphi$$

et comme les équations sont valables pour tout φ

$$\Theta\exp\left(\frac{\mathrm{i}}{\hbar}\,Ht\right) = \exp\left(-\frac{\mathrm{i}}{\hbar}\,Ht\right)\Theta \tag{A.18}$$

Si Θ était unitaire, cela impliquerait que

$$\Theta H = -H\Theta$$

et à tout vecteur propre φ_E de H d'énergie E correspondrait un vecteur propre $\Theta\varphi_E$ avec une énergie $-E$. Dans ces conditions, l'énergie ne serait pas bornée inférieurement et il existerait une instabilité fondamentale. Si au contraire Θ est antiunitaire, grâce à

$$\Theta \mathrm{i} H = -\mathrm{i}\Theta H$$

l'équation (A.18) implique

$$\Theta\, H = H\, \Theta \quad \text{ou} \quad \Theta\, H\, \Theta^{-1} = H \tag{A.19}$$

Cette dernière équation traduit l'invariance de H par renversement du sens du temps. Cependant, contrairement à l'opérateur parité Π, Θ ne conduit pas à une grandeur conservée, car l'équation (7.17) implique que l'opérateur A soit hermitien, ce qui n'est pas le cas de Θ. On sait aujourd'hui que toutes les interactions fondamentales de la physique sont invariantes par renversement du temps, sauf une interaction extrêmement faible, dont on ne voit les effets que dans le système de mésons $K^0 - \overline{K^0}$ (exercice 4.4.10), et que l'on a aussi observée tout récemment dans le système de mésons B, formés d'un quark ordinaire et d'un antiquark b (ou l'inverse).

Un double renversement du temps n'a évidemment aucun effet, et l'état $\Theta^2\varphi$ est équivalent à φ, $\Theta^2 = cI$, où c est un facteur de phase. La chaîne d'égalités

$$(\Theta\varphi_a, \varphi_b) = (\Theta\varphi_b, \Theta^2\varphi_a) = c(\Theta\varphi_b, \varphi_a) = c(\Theta\varphi_a, \Theta^2\varphi_b) = c^2(\Theta\varphi_a, \varphi_b)$$

montre que $c^2 = 1$, et donc $c = \pm 1$. Dans le cas où $c = -1$, le choix $\varphi_a = \varphi_b$ dans l'équation précédente entraîne

$$(\Theta\varphi_a, \varphi_a) = 0 \tag{A.20}$$

Si $c = -1$ et que H est invariant par renversement du temps, les états propres de H peuvent être rangés en paires d'états dégénérés par renversement du temps. Soit en effet φ, un vecteur propre de H : $H\varphi = E\varphi$. Alors,

$$H(\Theta\varphi) = \Theta(H\varphi) = E(\Theta\varphi)$$

et $\Theta\varphi$ est vecteur propre de H avec la valeur propre E : si $(\Theta\varphi, \varphi) = 0$, il existe (au moins) deux états propres de H avec la valeur propre E. Cette propriété est appelée *dégénérescence de Kramers*.

Compte tenu des propriétés de transformation de $\vec{J}$ (A.17), on doit avoir

$$\Theta|jm\rangle = \mathrm{e}^{\,\mathrm{i}\alpha}\,(-1)^{j-m}|j, -m\rangle \tag{A.21}$$

où, par application de J_+ et de J_-, on montre que α peut dépendre de j, mais non de m. On en déduit, en utilisant l'antilinéarité de Θ

$$\Theta^2|jm\rangle = (-1)^{2j}|jm\rangle \tag{A.22}$$

et $\Theta^2 = I$ si j est entier, $\Theta^2 = -I$ si j est demi-entier. La dégénérescence de Kramers entraîne alors qu'un système avec un nombre impair d'électrons possède des niveaux d'énergie doublement dégénérés en l'absence de champ magnétique. La présence d'un champ magnétique brise l'invariance par renversement du sens du temps, car pour respecter cette invariance, il faudrait inverser le sens des courants qui produisent ce champ : la raison pour laquelle l'effet Zeeman lève complètement la dégénérescence des niveaux est que le champ magnétique brise l'invariance par renversement du temps.

Pour un élément de matrice T (§ 13.5.4), $T_{ba} = T_{a\to b}$, l'invariance par renversement du temps implique la relation

$$T_{a\to b} = T_{\Theta b\to\Theta a} \tag{A.23}$$

où Θa (Θb), est l'état obtenu à partir de a (b) par renversement du temps, en inversant toutes les impulsions et tous les moments angulaires. Nous en déduisons par exemple la relation pour un élément de matrice T dans le cas de la diffusion de particules de spin zéro

$$T(\vec{k}_b, \vec{k}_a) = T(-\vec{k}_a, -\vec{k}_b)$$

Plus généralement, pour une réaction où les particules incidentes ont des impulsions $(\vec{p}_1, \vec{p}_2)$ et des projections de leurs spins (m_1, m_2), tandis que les particules finales sont caractérisées par $(\vec{p}_3, \vec{p}_4)$ et (m_3, m_4), on obtient

$$T_{m_3 m_4; m_1, m_2}(\vec{p}_1 + \vec{p}_2 \to \vec{p}_3 + \vec{p}_4) = T_{-m_1 -m_2; -m_3, -m_4}(-\vec{p}_3 - \vec{p}_4 \to -\vec{p}_1 - \vec{p}_2)$$

Pour une particule sans spin, l'opération de renversement du temps est simplement la conjugaison complexe. En effet, si $\psi(\vec{r}, t)$ vérifie l'équation de Schrödinger

$$\mathrm{i}\hbar \frac{\partial \psi(\vec{r}, t)}{\partial t} = -\frac{\hbar^2}{2m}\, \nabla^2 \psi(\vec{r}, t) + V(\vec{r})\psi(\vec{r}, t)$$

la fonction $[\Theta\psi](\vec{r}, t) = \psi^*(\vec{r}, -t)$ vérifie

$$\mathrm{i}\hbar \frac{\partial \psi^*(\vec{r}, -t)}{\partial t} = -\frac{\hbar^2}{2m}\, \nabla^2 \psi^*(\vec{r}, -t) + V(\vec{r})\psi^*(\vec{r}, -t)$$

pourvu que le potentiel $V(\vec{r})$ soit réel. Cette propriété a été utilisée au § 8.3.4 pour restreindre la forme de la matrice de passage M et de la matrice S.

Dans le cas général, on peut écrire $\Theta = UK$, où K est la conjugaison complexe

$$[K\psi](\vec{r}, t) = \psi^*(\vec{r}, -t) \tag{A.24}$$

et U un opérateur unitaire. L'action de K, par exemple, dans le cas d'une onde plane s'écrit

$$K\left[\mathrm{e}^{\,\mathrm{i}(\vec{k}\cdot\vec{r} - \omega t)}\right] = \mathrm{e}^{-\mathrm{i}(\vec{k}\cdot\vec{r} + \omega t)} = \mathrm{e}^{-\mathrm{i}\vec{k}\cdot\vec{r}}\, \mathrm{e}^{-\mathrm{i}\omega t}$$

ce qui correspond bien à une onde plane de vecteur d'onde $-\vec{k}$ et d'énergie positive $+\hbar\omega$. Pour une particule de spin 1/2, $U = \mathrm{i}\sigma_2$

$$\mathrm{i}\sigma_2|+\rangle = -|-\rangle \qquad \mathrm{i}\sigma_2|-\rangle = |+\rangle \tag{A.25}$$

en accord avec (A.21) si l'on choisit $\mathrm{e}^{\mathrm{i}\alpha} = -1$. On note que $(\mathrm{i}\sigma_2)^2 = -I$, également en accord avec (A.22) pour $j = 1/2$.

Comme dernier exemple, examinons l'impact de l'invariance par renversement du temps sur le moment dipolaire électrique du neutron. Comme l'opérateur moment dipolaire $\vec{D}$ est impair dans l'opération parité

$$\Pi\,\vec{D}\,\Pi^{-1} = -\vec{D}$$

le moment dipolaire[5] d'une particule est nul si cette particule a une parité déterminée, ce qui sera le cas si ses interactions conservent la parité. C'est pourquoi les atomes dans leur état fondamental n'ont pas de moment dipolaire permanent. Cependant, la parité n'est pas conservée dans les interactions faibles, et cela peut *a priori* rétablir la possibilité d'un moment dipolaire. En fait, il est de plus nécessaire que l'invariance par renversement du temps soit violée. En effet, le seul vecteur à notre disposition est le spin $\hbar\vec{\sigma}/2$ du neutron, et on doit avoir $\vec{D} = \lambda\vec{\sigma}$, où λ est une constante ; on remarque que $\lambda \neq 0$ implique la violation de la parité, car $\vec{D}$ est un vecteur et $\vec{\sigma}$ un pseudo-vecteur. Le couplage $\vec{D}\cdot\vec{E}$ du dipôle avec un champ électrique est impair par renversement du temps et doit s'annuler si l'invariance par renversement du temps est valide, car d'après (A.17), $\vec{\sigma}$ est impair et $\vec{E}$ pair : dans un renversement du temps, les charges sont inchangées (au contraire des courants, qui, comme on l'a vu, sont inversés). Si l'on envoie un neutron possédant un moment dipolaire électrique dans un champ électrique et un champ magnétique inhomogènes et constants et que l'on inverse à $t = 0$ la vitesse du neutron et les courants créant le champ magnétique, alors, à la différence de la figure A.1, le neutron ne "remontera pas sa trajectoire".

Essayons d'estimer le moment dipolaire du neutron par un argument dimensionnel. Ce moment dipolaire doit faire intervenir les interactions faibles, et donc la constante de Fermi G_F (exercice 13.6.6), ou plus précisément la combinaison $G_F/(\hbar c)^3$, et un paramètre sans dimension ε mesurant l'importance de la violation de l'invariance par renversement du temps, dont l'ordre de grandeur peut être estimé à environ 10^{-3} à partir de l'étude des mésons K neutres. On dispose en plus d'une masse, la masse du neutron $m_n \simeq 1\,\mathrm{GeV/c^2}$, et la seule solution possible est par analyse dimensionnelle

$$d \sim q_e\,\frac{G_F}{(\hbar c)^3}\,\varepsilon\,m_n(\hbar c^3)$$

5. Pour que le moment dipolaire électrique d'une particule puisse être non nul, il est impératif que son moment angulaire soit différent de zéro : dans le cas contraire, l'invariance par rotation est incompatible avec l'existence d'un moment dipolaire.

Il est commode d'utiliser un système d'unités où $\hbar = c = 1$ (exercice 13.6.1), 200 MeV $\simeq$ 1 fm^{-1}, ou 1 fm $\simeq 5\,\text{GeV}^{-1}$

$$d \sim q_e \times 10^{-5} \times 10^{-3} \times 1 = q_e \times 10^{-8}\,\text{GeV}^{-1} \sim q_e \times 10^{-9}\,\text{fm} = q_e \times 10^{-24}\,\text{m}$$

Les mesures les plus précises du moment dipolaire du neutron ont été effectuées au réacteur de recherches de l'Institut Laue-Langevin à Grenoble et donnent la borne supérieure

$$d \lesssim q_e \times 10^{-27}\,\text{m}$$

qui est largement inférieure à notre estimation naïve ! En fait, à cause d'une propriété technique du modèle standard[6], le moment dipolaire du neutron doit être proportionnel à G_F^2

$$d \sim G_F^2\,\varepsilon\, m_n^3(\hbar c^7) \simeq q_e \times 10^{-29}\,\text{m}$$

Les estimations théoriques du moment dipolaire du neutron ne sont pas très précises et varient aux alentours de $q_e \times 10^{-32}$ m : notre estimation est réduite par un facteur $\sim 10^{-3}$ car les calculs perturbatifs du modèle standard conduisent à un facteur numérique multiplicatif $\pi^{-4} \simeq 10^{-2}$ et suggèrent de prendre une masse caractéristique de l'ordre de 0.3 GeV au lieu de m_n. La violation de Θ a été observée en 2012 : Lees *et al.* [2012], Schwarzschild [2012].

B Méthode de Wigner et Weisskopf

La déduction de la règle d'or de Fermi du § 8.5.3 est limitée à des temps suffisamment courts, $t \ll \tau_2$, et il n'est pas possible de justifier avec les seuls arguments du § 8.5.3 la loi de décroissance exponentielle (8.153). Une méthode due à Wigner et Weisskopf[7] permet de justifier cette loi pour des temps longs à l'aide d'un autre schéma d'approximations. Considérons la situation suivante : un état d'un système isolé a d'énergie E_a, se désintègre vers un continuum d'états b, d'énergie E_b. Des exemples d'une telle situation sont la désexcitation d'un état excité d'un atome, d'une molécule, d'un noyau atomique... avec émission d'un photon, ou la désintégration d'une particule élémentaire. Les états d'énergie E_a et E_b sont états propres d'un hamiltonien $H^{(0)}$

$$H^{(0)}|a\rangle = E_a|a\rangle \qquad H^{(0)}|b\rangle = E_b|b\rangle \tag{B.1}$$

et une perturbation W indépendante du temps est responsable de la transition $a \to b$; dans le cas de l'émission spontanée d'un photon, W est donné par (17.122). Les états a et b ne sont *pas* des états stationnaires du hamiltonien

6. Voir par exemple Donoghue *et al.* [1992], chapitre IX.

7. La méthode de Wigner et Weisskopf est décrite par Cohen-Tannoudji *et al.* [1973], complément D_{XIII}, ou Basdevant et Dalibard [2001], chapitre 17 ; un traitement détaillé et rigoureux est donné par Messiah [1959], chapitre XXI.

total indépendant du temps $H = H^{(0)} + W$. Nous pouvons supposer que les éléments de matrice diagonaux de W sont nuls[8] : $W_{aa} = W_{bb} = 0$ et nous notons $|\psi(t)\rangle$ le vecteur d'état du système, dont l'état initial est $|\psi(t=0)\rangle = |a\rangle$. Décomposons l'état $|\psi(t)\rangle$ sur les états $|a\rangle$ et $|b\rangle$ en faisant intervenir la densité d'états $\mathcal{D}(E_b)$

$$|\psi(t)\rangle = \gamma_a(t)\mathrm{e}^{-\mathrm{i}E_a t/\hbar}|a\rangle + \int \mathrm{d}E_b\,\mathcal{D}(E_b)\,\gamma_b(t)\,\mathrm{e}^{-\mathrm{i}E_b t/\hbar} \tag{B.2}$$

L'équation de Schrödinger appliquée sur la décomposition (B.2)

$$\mathrm{i}\hbar\frac{\mathrm{d}|\psi(t)\rangle}{\mathrm{d}t} = \left(H^{(0)} + W\right)|\psi(t)$$

conduit au système d'équations différentielles (*cf.* (8.144))

$$\mathrm{i}\hbar\dot{\gamma}_a(t) = \int \mathrm{e}^{\,\mathrm{i}\omega_{ab}t}W_{ab}\,\gamma_b(t)\mathcal{D}(E_b)\mathrm{d}E_b \tag{B.3}$$

$$\mathrm{i}\hbar\dot{\gamma}_b(t) = \mathrm{e}^{-\mathrm{i}\omega_{ab}t}\,W^*_{ab}\,\gamma_a(t) \tag{B.4}$$

avec $\omega_{ab} = (E_a - E_b)/\hbar$. Nous savons empiriquement que $|\gamma_a(t)|^2$ est donné par une loi exponentielle

$$|\gamma_a(t)|^2 = \mathrm{e}^{-\Gamma t} \tag{B.5}$$

ce qui suggère d'essayer un comportement

$$\gamma_a(t) = \exp\left(-\frac{\mathrm{i}\delta}{2}\,t\right) \qquad \delta = \delta_1 - \mathrm{i}\Gamma \tag{B.6}$$

où δ_1 est réel. La substitution de (B.6) dans (B.4) avec les conditions initiales $\gamma_b(t=0) = 0$ donne par intégration sur t

$$\gamma_b(t) = \frac{W^*_{ab}}{\hbar\,[\omega_{ab} + \delta/2]}\,\left[\exp\left(-\mathrm{i}(\omega_{ab} + \delta/2)t\right) - 1\right] \tag{B.7}$$

Pour des temps longs, $t \gg \Gamma^{-1}$, l'exponentielle dans (B.7) tend rapidement vers zéro et

$$\lim_{t\to\infty}\,|\gamma_b(t)|^2 = \frac{|W_{ab}|^2}{\hbar^2\left[(\omega_{ab} + \delta_1/2)^2 + \Gamma^2/4\right]} \tag{B.8}$$

Pour vérifier la cohérence de notre hypothèse de départ avec les équations d'évolution (B.3) et (B.4), il faut reporter (B.6) dans (B.3), ce qui donne

$$\frac{\hbar\delta}{2} = \int \mathrm{d}E_b\,\mathcal{D}(E_b)\,|W_{ab}|^2\,\frac{1 - \exp(\mathrm{i}[\omega_{ab} + \delta/2]t)}{\hbar(\omega_{ab} + \delta/2)} \tag{B.9}$$

8. Si tel n'était pas le cas, on pourrait toujours redéfinir $H^{(0)}$

$$H^{(0)} \to H^{(0)\prime} = H^{(0)} + |a\rangle\,W_{aa}\,\langle a| + \int \mathrm{d}E_b\,\mathcal{D}(E_b)\,|b\rangle\,W_{bb}\,\langle b|$$

La constante δ doit être solution de l'équation intégrale (B.9). Pour fixer les idées, examinons la transition d'un état excité i d'un atome d'énergie E_i vers l'état fondamental f d'énergie E_f de cet atome, avec émission d'un photon d'énergie $\hbar\omega$. On peut, à une excellente approximation, négliger l'énergie cinétique de recul de l'atome final dont l'énergie est simplement E_f avec comme choix de référentiel celui où l'atome est au repos dans son état initial (*cf.* la discussion du § 15.3.3). La densité d'états finaux à utiliser dans (B.9) est celle (17.126) du photon. En résumé, nous avons $|a\rangle = |i\rangle$ et $|b\rangle$ = atome dans l'état $f+$ photon $= |f \otimes \vec{k}s\rangle$, ainsi que la conservation de l'énergie

$$\hbar\omega_{ab} = E_a - E_b = E_i - (E_f + \hbar\omega) = \hbar(\omega_0 - \omega) \tag{B.10}$$

avec $\hbar\omega_0 = E_i - E_f$. L'équation (B.9) devient, en choisissant ω comme variable d'intégration au lieu de E_f, $\mathrm{d}E_b = \hbar \mathrm{d}\omega$

$$\frac{\hbar\delta}{2} = \int_0^\infty \mathrm{d}\omega\, \mathcal{D}(\omega)\, |W_{ab}(\omega)|^2\, \frac{1 - \exp(\mathrm{i}[\omega_0 - \omega + \delta/2]t)}{\omega_0 - \omega + \delta/2} \tag{B.11}$$

Nous sommes intéressés par le comportement de cette équation aux temps longs, et nous avons besoin du comportement pour $t \to \infty$ de la fonction $f(t,x)$ considérée comme distribution

$$f(t,x) = \frac{1 - \mathrm{e}^{\,itx}}{x}$$

Lorsque x est réel, sa transformée de Fourier est

$$\tilde{f}(t,u) = \mathrm{i}[\theta(u) - \theta(t+u)] \tag{B.12}$$

car

$$-\mathrm{i}\int_{-t}^{0} \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}ux} = \frac{1 - \mathrm{e}^{\,itx}}{x}$$

et la limite $t \to \infty$ de $\tilde{f}(t,u)$ est simplement $-\mathrm{i}\theta(-u)$, ce qui donne pour $f(t,x)$

$$\lim_{t\to\infty} f(t,x) = \lim_{\eta\to 0+} \frac{1}{x + \mathrm{i}\eta} = \mathbf{P}\,\frac{1}{x} - \mathrm{i}\pi\delta(x) \tag{B.13}$$

Ce résultat est aussi valable lorsque x a une petite partie imaginaire, $x = \mathrm{Re}\, x \pm \mathrm{i}\eta$, $\eta \to 0^+$. Il suffit pour le voir d'intégrer sur x dans le plan complexe, en complétant le contour d'intégration par un demi-cercle dont le rayon tend vers l'infini. Si δ_1 et Γ sont petits par rapport aux intervalles caractéristiques de variation des fonctions $\mathcal{D}(\omega)$ et $|W_{ab}(\omega)|^2$, on peut reporter (B.13) dans (B.11) et déterminer la valeur de δ

$$\delta = \frac{2}{\hbar}\,\mathbf{P}\int_{-\infty}^{\infty} \mathrm{d}\omega\, \frac{\mathcal{D}(\omega)|W_{ab}(\omega)|^2}{\omega_0 - \omega} - \frac{2\mathrm{i}\pi}{\hbar}\,\mathcal{D}(\omega_0)|W_{ab}(\omega_0)|^2 \tag{B.14}$$

Le deuxième terme du membre de droite de (B.14) confirme que $\Gamma = -\mathrm{i}\,\mathrm{Im}\,\delta$ est bien donné par la règle d'or de Fermi

$$\Gamma = \frac{2\pi}{\hbar}\,\mathcal{D}(\omega_0)|W_{ab}(\omega_0)|^2 \tag{B.15}$$

tandis que le premier terme correspond au déplacement du niveau d'énergie

$$\mathrm{Re}\,\delta = \delta_1 = \frac{2}{\hbar}\,\mathbf{P}\int_{-\infty}^{\infty}\mathrm{d}\omega\,\frac{\mathcal{D}(\omega)|W_{ab}(\omega)|^2}{\omega_0-\omega} \tag{B.16}$$

Ce déplacement aurait pu être obtenu par un calcul au second ordre en théorie des perturbations indépendantes du temps ; il est nul au premier ordre d'après notre hypothèse $W_{aa} = W_{bb} = 0$. On peut absorber δ_1 dans une redéfinition de ω_0 : $\omega_0 \to \omega_0 + \delta_1$ et d'après (B.8), la probabilité d'observer un photon de fréquence ω est

$$\mathsf{p}(\omega)\,\mathrm{d}\omega \simeq \frac{\mathcal{D}(\omega_0)|W_{ab}(\omega_0)|^2}{\hbar\big[(\omega-\omega_0)^2+\Gamma^2/4\big]}\,\mathrm{d}\omega \tag{B.17}$$

Cette probabilité est correctement normalisée à l'unité

$$\int_{-\infty}^{+\infty}\mathsf{p}(\omega)\,\mathrm{d}\omega = 1 \quad \text{car} \quad \int_{-\infty}^{+\infty}\frac{\mathrm{d}x}{x^2+a^2} = \frac{\pi}{a}$$

compte tenu de la valeur (B.15) de Γ. La courbe représentative de $\mathsf{p}(\omega)$ est une lorentzienne (aussi appelée courbe de Breit-Wigner)

$$\mathsf{p}(\omega) = \frac{\mathcal{D}(\omega_0)|W_{ab}(\omega_0)|^2}{\hbar\big[(\omega-\omega_0)^2+\Gamma^2/4\big]} \tag{B.18}$$

La fréquence du photon final n'a pas une valeur bien déterminée : elle présente une “dispersion"[9] $\Delta\omega = \Gamma$, définie comme la largeur à mi-hauteur de la courbe $\mathsf{p}(\omega)$

$$\mathsf{p}\left(\omega_0 \pm \frac{1}{2}\,\Delta\omega\right) = \frac{1}{2}\mathsf{p}(\omega=\omega_0)$$

En d'autres termes, le spectre de fréquences du photon émis n'est pas monochromatique. La quantité Γ est appelée *largeur de raie* ou parfois largeur naturelle de raie, étant donné qu'il existe d'autres causes d'élargissement comme l'effet Doppler ou les collisions ; en raison de (B.5), la vie moyenne de l'état excité est l'inverse de la largeur de raie, $\tau = 1/\Gamma$. La dispersion en énergie du photon final montre, par conservation de l'énergie, que l'énergie de l'état excité présente une dispersion $\Delta E = \hbar\Gamma$, et nous en déduisons la relation (4.29) entre la vie moyenne et la dispersion sur l'énergie

9. Dispersion est à mettre entre guillemets car l'intégrale

$$\int_0^{\infty}\mathrm{d}\omega(\omega-\omega_0)^2\mathsf{p}(\omega)$$

est divergente, et on ne peut pas définir une dispersion au sens strict du terme.

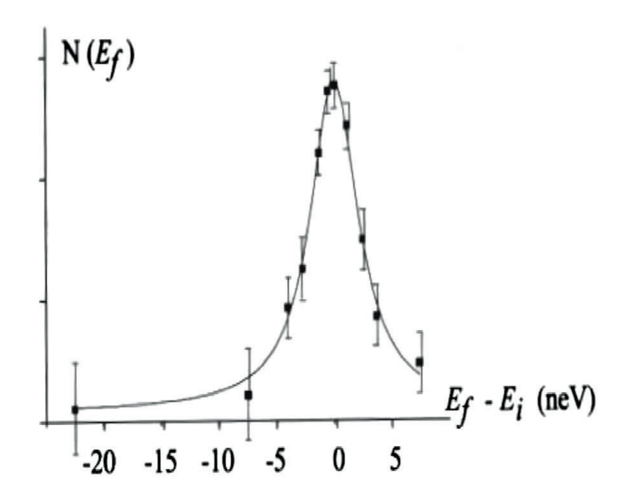


FIG. B.2 – Spectre des photons de désintégration ${}^{57}Fe^* \to {}^{57}Fe+$ photon. D'après Basdevant et Dalibard [2001].

$$\tau\,\Delta E = \hbar \qquad \text{(B.19)}$$

La conservation de l'énergie implique que l'énergie de l'état excité n'a pas une valeur précise, mais présente une dispersion $\Delta E \simeq \hbar\Gamma$ autour de sa valeur centrale E_i. Cependant, il n'y a pas de limitation de principe à la précision avec laquelle on peut mesurer cette valeur centrale. La figure B.2 montre la détermination expérimentale de $\mathsf{p}(\omega)$ pour la désintégration d'un niveau excité du ${}^{57}\mathrm{Fe}^*$

$$ {}^{57}\mathrm{Fe}^* \to {}^{57}\mathrm{Fe} + \text{ photon } (14\,\mathrm{keV})$$

dont la vie moyenne est $\tau \simeq 1.4 \times 10^{-7}\,$s.

C Constantes physiques

vitesse de la lumière dans le vide	:	$c = 3.00\,\mathrm{m.s}^{-}1$
constante de Planck	:	$h = 6.63 \times\ 10^{-34}\mathrm{J.s}$
constante de Planck divisée par 2π	:	$\hbar = 1.055 \times\ 10^{-34}\mathrm{J.s}$
charge de l'électron ($q_e < 0$)	:	$\lvert q_e \rvert = 1.602 \times\ 10^{-19}\ \mathrm{Cb}$
masse de l'électron	:	$m_e = 9.11 \times\ 10^{-31}\,\mathrm{kg} = 0.511\,\mathrm{MeV.c}^{-2}$
masse du proton	:	$m_p = 1.673 \times\ 10^{-27}\,\mathrm{kg} = 938.3\,\mathrm{MeV.c}^{-2}$
masse du neutron	:	$m_n = 1.675 \times\ 10^{-27}\,\mathrm{kg} = 939.5\,\mathrm{MeV.c}^{-2}$
constante de Rydberg	:	$R_\infty = \frac{1}{2}\alpha^2 m_e c^2 = 13.61\,\mathrm{eV}$
magnéton de Bohr	:	$\mu_\mathrm{B} = \frac{\lvert q_e \rvert \hbar}{2m_e} = 5.79 \times 10^{-5}\,\mathrm{eV.T}^{-1}$
magnéton de Bohr nucléaire	:	$\mu_N = \frac{\lvert q_e \rvert \hbar}{2m_p} = 3.15 \times 10^{-8}\,\mathrm{eV.T}^{-1}$
constante de Boltzmann	:	$k_\mathrm{B} = 1.38 \times 10^{-23}\,\mathrm{J.K}^{-1}$
constante de gravitation	:	$G = 6.67 \times 10^{-11}\,\mathrm{N.m}^2.\mathrm{kg}^{-2}$
conversion énergie-longueur d'onde	:	$E = 1\,\mathrm{eV} \leftrightarrow \lambda = 1.24\,\mu\mathrm{m}$
conversion énergie-fréquence	:	$E = 1\,\mathrm{eV} \leftrightarrow \nu = 2.42 \times 10^{14}\,\mathrm{Hz}$
conversion énergie-température	:	$E = 1\,\mathrm{eV} \leftrightarrow T = 11600\,\mathrm{K}$

Références

1. Abers [2004] : E. Abers, *Quantum mechanics*, Pearson Education, New Jersey.
2. Abramowitz et Stegun [1964] : M. Abramowitz and I. Stegun eds., *Handbook of mathematical functions*, Dover, New-York.
3. Ahmad *et al.* [2001] : Q. Ahmad *et al.*, (SNO Collaboration), Interactions produced by B^8 solar neutrinos at the Sudbury Neutrino Observatory, *Phys.Rev. Lett.* **87**, 071301.
4. Akitt [1992] : J.W. Akitt, *NMR chemistry, an introduction to modern NMR spectroscopy*, Chapman&Hall, New-York.
5. Anandan [1992] : J. Anandan, The geometric phase, *Nature* **360**, 307.
6. Arndt *et al.* [1999] : M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. van der Zouw and A. Zeilinger, Wave-particle duality of C_{60} molecules, *Nature* **401**, 680 (1999).
7. Arndt *et al.* [2005] : M. Arndt, K. Hornberger and A. Zeilinger, Probing the limits of the quantum world, *Physics World* **18**, March 2005, pp. 35–40.
8. Aspect [1999] : A. Aspect, Bell's inequalities : more ideal than ever, *Nature* **398**, 189-190.
9. Aspect *et al.* [1982] : A. Aspect, P. Grangier, and G. Roger, Experimental realization of Einstein-Podolsky-Rosen gedanken experiment : a new violation of Bell's inequalities, *Phys. Rev. Lett.* **49**, 91 (1982) ; A. Aspect, J. Dalibard, and G. Roger, Experimental test of Bell's inequalities using time varying analyzers, *Phys. Rev. Lett.* **49** 1804 (1982).
10. Aspect *et al.* [1989] : A. Aspect, P. Grangier et G. Roger, Dualité onde-corpuscule pour un photon unique, *J. Optics (Paris)* **20**, 119.
11. Badurek *et al.* [1985] : G. Badurek, H. Rauch and D. Tuppinger, Neutron interferometric double-resonance experiment, *Phys. Rev.* **A34**, 2600-2608.
12. Balian [1991] : R. Balian, *From Microphysics to Macrophysics*, Springer, Berlin.
13. Ballentine [1990] : L. Ballentine, Limitations of the projection postulate, *Found. Phys.* **20**, 1329.
14. Ballentine [1998] : L. Ballentine, *Quantum Mechanics*, World Scientific.

15. Basdevant et Dalibard [2001] : J-L Basdevant et J. Dalibard, *Mécanique quantique*, Éditions de l'École Polytechnique, Palaiseau.
16. Baym [1969] : G. Baym, *Lectures on Quantum Mechanics*, Benjamin, Reading.
17. Baym [1998] : G. Baym, The physics of Hanbury Brown-Twiss intensity interferometry : from stars to nuclear collisions, *Acta Physica Polonica* **B29**, 1839.
18. Bell [1990] : J. S. Bell Against measurement, *Physics World* **33**, August, p. 33.
19. Bell [2004] : J. S. Bell *Speakable and unspeakable in quantum mechanics*, Cambridge University Press, Cambridge.
20. Bennett *et al.* [1992] : C. Bennett, G. Brassard and A. Ekert, Quantum cryptography, *Scientific American* **267**, October p. 26.
21. Berry [1991] : M. Berry, Some Quantum to Classical Asymptotics, Les Houches 1989, *Chaos et Physique Quantique*, Elsevier.
22. Berry et Robbins [1997] : M. Berry and J. Robbins, Indistiguishability for quantum particles : spin, statistics and the geometric phase, *Proc. Roy. Soc. London* **A453**, 1771.
23. Bethe et Morrison [1956] : H. Bethe et Ph. Morrison, *Elementary Nuclear Theory*, John Wiley.
24. Blatt [2004] : R. Blatt, Quantum information processing in ion traps, in Les Houches Summer School 2003, *Quantum entanglement and information processing*, D. Estève, J.-M. Raimond, and M. Brune eds., Elsevier, Amsterdam, p. 223-260.
25. Bohr [1935] : N. Bohr, Can quantum mechanical description of physical reality be considered complete ? *Phys. Rev.* **48**, 696.
26. Braginsky *et al.* [1980] : V. Braginsky, Y. Vorontsov and K. Thorne, Quantum non demolition measurements, *Science* **209**, 547.
27. Bransden et Joachain [1983] : B. Bransden and C. Joachain, *Physics of Atoms and Molecules*, Longman Scientific and Technical, Harlow (1983)
28. Breitenaker et Treps [2010] : F. Breitenaker et N. Treps eds, *Le laser* EDP Sciences, Paris.
29. Breuer et Petruccione [2002] : H. Breuer and F. Petruccione, *The theory of open quantum systems*, Oxford University Press, Oxford.
30. Brown et Holland [1999] : H. Brown and P. Holland, The Galilean covariance of quantum mechanics in the case of external fields, *Am. Journ. Phys.* **67**, 204.
31. Brune *et al.* [1996a] : M. Brune, F. Schmidt-Kaler, A. Maali, J. Dreyer, E. Hagley, J-M Raimond and S. Haroche, Quantum Rabi oscillations : a direct test of field quantization, *Phys. Rev. Lett.* **76**, 1800.
32. Brune *et al.* [1996b] : M. Brune, E. Hagley, J. Dreyer, X. Maître, A. Maali, C. Wunderlich, J-M Raimond and S. Haroche, Observing the progressive decoherence of the "meter" in a quantum measurement, *Phys. Rev. Lett.* **77**, 4887 (1996).
33. Buhrman *et al.* [2010] : H. Buhrman, R. Cleve, S. Massar and R. de Wolf, Nonlocality and communication complexity, *Rev. Mod. Phys.* **82**, 665.

34. Buzek et Hillery [1996] : V. Buzek and M. Hillery, Quantum copying : beyond the no-cloning theorem, *Phys. Rev.* **A54**, 1844.
35. Caldeira et Leggett [1983] : A. Caldeira and A. Leggett, Path integral approach to quantum Brownian motion, *Ann. Phys.* **149**, 374.
36. Carmichael [1993] : H. Carmichael, *An open system approach to quantum optics*, Springer-Verlag, Berlin.
37. Castin [2001] : Y. Castin, Bose-Einstein condensates in atomic gases : simple theoretical results, dans *Coherent atomic matter waves*, École d'été des Houches, R. Kaiser, C. Westbrook, and F. David eds., EDP Sciences/Springer-Verlag.
38. Caves *et al.* [1980] : C. Caves *et al.*, On the measurement of a weak classical force coupled to a quantum mechanical oscillator, *Rev. Mod. Phys.* **52**, 341.
39. Child [1991] : M. Child, *Semiclassical mechanics with molecular applications*, Clarendon Press, Oxford.
40. Clauser [1976] : J. Clauser, Experimental investigation of a polarization correlation anomaly, *Phys. Rev. Lett.* **36**, 1223.
41. Cohen-Tannoudji [1989] : C. Cohen-Tannoudji, Cohérences quantiques et dissipation, Cours au Collège de France 1988/89, disponible sur le site web www.lkb.ens.fr.
42. Cohen-Tannoudji [1990] : C. Cohen-Tannoudji, Théorie de la mesure, Cours au Collège de France 1989/90, disponible sur le site web www.lkb.ens.fr.
43. Cohen-Tannnoudji [1992] : C. Cohen-Tannnoudji, *Atomic Motion in Laser Light*, dans *Optical Coherence and Quantum Optics*, École des Houches, North Holland.
44. Cohen-Tannoudji [1999] : C. Cohen-Tannoudji, *Condensats de Bose-Einstein atomiques*, Cours au Collège de France 1998/1999, disponible sur le site web www.lkb.ens.fr.
45. Cohen-Tannoudji *et al.* [1973] : C. Cohen-Tannoudji, B. Diu et F. Laloë, *Mécanique quantique*, Hermann, Paris.
46. Cohen-Tannoudji *et al.* [1988] : C. Cohen-Tannoudji, J. Dupont-Roc et G. Grynberg, *Processus d'interaction entre photons et atomes*, InterEditions/Editions du CNRS, Paris.
47. Cohen-Tannoudji *et al.* [2005] : C. Cohen-Tannoudji, J. Dalibard et F. Laloë, La condensation de Bose-Einstein dans les gaz, dans *Einstein aujourd'hui*, p. 87, M. Le Bellac et M. Leduc eds., EDP Sciences/Éditions du CNRS, Paris.
48. Colella *et al.* [1975] : R. Colella, A. Overhauser and S. Werner, Observation of gravitationally induced quantum interference, *Phys. Rev. Lett.* **34**, 1472.
49. Comtet *et al.* [1992] : A. Comtet, J. Mc Cabe et S. Ouvry, Vers de nouvelles statistique quantiques, *Images de la Physique*, CNRS, p. 21.
50. Conway et Kochen [2009] ; J. Conway and S. Kochen, The strong free will theorem, *Notices of the AMS* **56**, 226.
51. Dalibard [2003] : J. Dalibard, *Cohérence quantique et dissipation*, disponible sur le site web www.lkb.ens.fr.
52. Darrigol [2005] : O. Darrigol, 1905 : un nouvel élan, dans *Einstein aujourd'hui*, p. 1, M. Le Bellac et M. Leduc eds., EDP Sciences/Éditions du CNRS, Paris.

53. Diu *et al.*[1990] : B. Diu, C. Guthman, D. Lederer et B. Roulet, *Physique statistique*, Hermann, Paris.
54. Donoghue *et al.* [1992] : J. Donoghue, E. Golowich and B. Holstein, *Dynamics of the standard model*, Cambridge University Press, Cambridge.
55. Dürr *et al.* [1998] : S. Dürr, T. Nonn, and G. Rempe, Origin of quantum-mechanical complementarity probed by a 'which way' experiment in an atom interferometer, *Nature* **395**, 33–37.
56. Eguchi *et al.* [2003] : K. Eguchi *et al.* (Kamland Collaboration), First results from Kamland : evidence for reactor neutrino disappearance, *Phys. Rev. Lett.* **90**, 021802.
57. Einstein *et al.* [1935] : A. Einstein, B. Podolsky and N. Rosen, Can quantum-mechanical description of physical reality be considered complete ?, *Phys. Rev.* **47**, 777–780.
58. Englert *et al.* [1991] : B. Englert, M. Scully and H. Walther, Quantum optical tests of complementarity, *Nature* **351**, 111.
59. Englert *et al.* [1995] : B. Englert, M. Scully and H. Walther, La dualité de la lumière et de la matière, *Pour la Science* **208**, 60.
60. d'Espagnat [1995] : B. d'Espagnat, *Le réel voilé*, Fayard, Paris.
61. Feynman *et al.* [1965] : R. Feynman, R. Leighton and M. Sands, *The Feynman Lectures on Physics*, Addison-Wesley, Reading.
62. Fine [1982] : A. Fine, Hidden variables, joint probability and the Bell inequalities, *Phys. Rev. Lett.* **48**, 291.
63. Foerster [1975] : D. Foerster, *Hydrodynamics fluctuations, broken symmetry and correlation functions*, Benjamin, New York.
64. Freedman et Clauser [1972] : S. Freedman and J. Clauser, Experimental test of local hidden variable theories, *Phys. Rev. Lett.*, **28**, 938.
65. Fukuda *et al.* [2001] : S. Fukuda *et al.* (SuperKamiokande Collaboration), Solar B^8 and high energy physics neutrino measurements from 1258 days of SuperKamiokande data, *Phys.Rev. Lett.* **86**, 5651.
66. Gasiorowicz [1966] : S. Gasiorowicz, *Elementary Particle Physics*, J. Wiley, New-York.
67. Gatz [1971] : C. Gatz, *Introduction to Quantum Chemistry*, C.E. Merrill, Columbia.
68. Gerry et Knight [2005] : C. Gerry and P. Knight, *Introductory Quantum Optics*, Cambridge University Press.
69. Ginsberg *et al.* [2007] : N. Ginsberg, S. Garner and L. Hau, Coherent control of optical information with matter wave dynamics, *Nature* **445**, 623.
70. Gisin [2012] : N. Gisin, *L'impensable hasard*, Odile Jacob, Paris.
71. Gisin *et al.* [2002] : N. Gisin, G. Ribordy, W. Tittel and H. Zbinden, Quantum cryptography, *Rev. Mod. Phys.* **74**, 145–195.
72. Gleyzes *et al.* [2007] : S. Gleyzes *et al.*, Quantum jumps of light recording the birth and death of a photon in a cavity, *Nature* **446**, 297.
73. Goodrich [1972] : F. Goodrich, *A Primer of Quantum Chemistry*, Wiley, New-York.

74. Gottfried et Yan [2033] ; K. Gottfried and T. Yan, *Quantum mechanics : fundamentals*, Springer, Berlin.
75. Gourgoulhon [2010] : E. Gourgoulhon, *Relativité restreinte*, EDP Sciences/CNRS Editions, Paris.
76. Grangier [2002] : Ph. Grangier, Single photons stick together, *Nature* **419**, 577.
77. Grynberg *et al.* [2010] : G. Grynberg, A. Aspect and C. Fabre, *Introduction to quantum optics*, Cambridge University Press, Cambridge.
78. Guerlin *et al.* [2007] : C. Guerlin *et al.*, Progressive field-state collapse and quantum non-demolition photon counting, *Nature* **448**, 889.
79. Gutzwiller [1990] : M. Gutzwiller, *Chaos in Classical and Quantum Systems*, Springer, Berlin.
80. Haroche et Raimond [2006] : S. Haroche and J.-M. Raimond, *Exploring the quantum*, Oxford University Press, Oxford.
81. Hasegawa it et al. [2006] : Y. Hasegawa, R. Loidl, G. Badurek, M. Baron and H. Rauch, Quantum contextuality in a single-neutron optical experiment, *Phys. Rev. Lett.* **97**, 230401.
82. Hecht [1987] : E. Hecht, *Optics*, Addison-Wesley, New-York.
83. Hey et Walters [2003] : T. Hey and P. Walters, *The New Quantum Universe*, Cambridge University Press, Cambridge.
84. Hillery *et al.* [1984] : M. Hillery, R. O'Connell, M. Scully and E. Wigner, Distribution functions in Physics : Fundamentals, *Phys. Reports* **106**, 121-167.
85. Holland [1993] : P. Holland, *The quantum theory of motion*, Cambridge University Press, Cambridge.
86. Hornberger *et al.* [2012] : K. Hornberger, S. Gerlich, P. Haslinger, S. Nimmrichter and M. Arndt, Quantum interference of clusters and molecules, *Rev. Mod. Phys.* **84**, 157.
87. Howard [2004] : D. Howard, Who invented the Copenhagen interpretation ?, *Phil. of Science* **71**, 669.
88. Huang [1963] : K. Huang, *Statistical Mechanics*, John Wiley, New-York.
89. Isham [1995] : C. Isham, *Lectures on Quantum Theory*, Imperial College Press, Londres.
90. Itzykson et Zuber [1980] : C. Itzykson et J-B Zuber, *Quantum Field Theory*, Mac Graw Hill, New-York.
91. Jackson [2001] : J.D. Jackson, *Électrodynamique classique*, Dunod, Paris.
92. Jacques *et al.* [2005] : V. Jacques, E. Wu, F. Grosshans, F. Treussart, Ph. Grangier, A. Aspect and J-F Roch, Single photon wave-front splitting interference, *Eur. J. Phys.* **D 35**, 561.
93. Jacques *et al.* [2007] : V. Jacques, E. Wu, F. Grosshans, F. Treussart, Ph. Grangier, A. Aspect and J-F Roch, Experimental realization of Wheeler's delayed choice experiment, *Science* **315**, 966.
94. Jauch [1968] : J. Jauch, *Foundations of Quantum Mechanics*, Addison Wesley, Reading.

95. Jacob [2002] : M. Jacob, Le modèle standard de la physique des particules, *Pour la Science* **300**, p. 58, octobre.
96. Jeltes *et al.* [2007] : T. Jeltes *et al.*, Comparison of the Hanbury Brown-Twiss effect for bosons and fermions, *Nature* **445**, 402.
97. Joos et Zeh [1985] : E. Joos and H. Zeh : The emergence of classical properties through interaction with the environment, *Zeit. Phys.* **B59**, 223.
98. Kaiser *et al.* [2012] : F. Kaiser, T. Coudreau, P. Milman, D. Ostrowsky and S. Tanzilli, Entanglement enabled delayed choice experiment, *Science* **338**, 637.
99. Kiesel *et al.* [2002] : H. Kiesel, A. Renz and F. Hasselbach, Observation of Hanbury Brown and Twiss anticorrrelations for free electrons, *Nature* **418**, 392.
100. Kittel [1970] : C. Kittel, *Introduction à la physique du solide*, Dunod, Paris.
101. Kragh [2000] : H. Kragh, Max Planck : the reluctant revolutionary, *Physics World* **13**(12) December.
102. Landau et Lifschitz [1966] : L. Landau et E. Lifchitz, *Mécanique quantique*, Éditions MIR, Moscou.
103. Laloë [2011] : F. Laloë, *Comprenons-nous vraiment la mécanique quantique ?*, EDP Sciences/CNRS Éditions, Paris.
104. Le Bellac [1988] : M. Le Bellac, *Des phénomènes critiques aux champs de jauge*, Interéditions/CNRS Éditions, Paris.
105. Le Bellac [2006] : M. Le Bellac, *A short introduction to quantum information and computation*, Cambridge University Press, Cambridge.
106. Le Bellac [2010] : M. Le Bellac, *Le monde quantique*, EDP Sciences, Paris.
107. Le Bellac et Lévy-Leblond [1973] : M. Le Bellac and J.-M. Lévy-Leblond, Galilean Electromagnetism, *Nuovo Cimento* **B14**, 217.
108. Le Bellac *et al.* [2004] : M. Le Bellac, F. Mortessagne and George Batrouni *Equilibrium and Non Equilibrium Statistical Thermodynamics*, Cambridge University Press, Cambridge.
109. Lees *et al.* [2012] : J. Lees *et al.* (Babar collaboration), Observation of time-reversal violation in the B^0 meson system, arxiv : 1207.5832.
110. Leggett [2001] : A. Leggett, Bose-Einstein condensation in the alkali gases : some fundamental concepts, *Rev. Mod. Phys.* **73**, 307.
111. Leggett [2002a] : A. Leggett, Qubits, cbits, quantum measurement and environment, in *Fundamentals of quantum information*, D. Heiss editor, Lecture notes in Physics **587**, p. 3, Springer-Verlag, Berlin.
112. Leggett [2002b] : A. Leggett, Testing the limits of quantum mechanics, *J. Phys. Cond. Mat.* **14**, R415.
113. Leggett [2005] : A. Leggett, The quantum measurement problem, *Science* **307**, 871.
114. Leggett [2006] : A. Leggett, *Quantum fluids*, Oxford University Press, Oxford.
115. Leggett [2008] : A. Leggett, Realism and the physical world, *Rep. Prog. Phys* **71**, 1.
116. Levitt [2001] : M. H. Levitt, *Spin Dynamics, Basics of Nuclear Magnetic Resonance*, John Wiley, New-York.

117. Lévy-Leblond [1967] : J.-M. Lévy-Leblond, Non relativistic particles and wave equations, *Comm. Math. Phys.* **6**, 286.
118. Lévy-Leblond [1999] : J.-M. Lévy-Leblond, Mots et maux de la physique quantique, *Bull. U. Phys.* **816**, 1129.
119. Lévy-Leblond et Balibar [1984] : J-M Lévy-Leblond et F. Balibar, *Quantique : Rudiments*, InterEditions, Paris.
120. Mandel and Wolf [1995] : L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, Cambridge University Press, Cambridge.
121. Matsukevitch *et al.* [2008] : D. Matsukevitch, P.Maunz, D. Moering, S. Olmscheck and C. Monroe, Bell inequality violation with two remote atomic qubits, *Phys. Rev. Lett.* **100**, 150404.
122. May et Cazabat [1996] : M. May et A-M Cazabat, *Optique*, Dunod, Paris.
123. Mazoyer [2002] : B. Mazoyer, L'imagerie cérébrale : IRM et TEP, *Pour la Science* **302**, décembre, p. 42.
124. Mermin [1993] : N. Mermin, Hidden variables and the two theorems of John Bell, *Rev. Mod. Phys.* **65**, 803.
125. Mermin [1998] : N. Mermin, What is Quantum Mechanics trying to teach us ?, *Am. Journ. Phys.* **66**, 753.
126. Mermin [2010] : N. Mermin, *Calculs et algorithmes quantiques : méthodes et exemples*, traduction de S. Tanzilli et O. Alibart, EDP Sciences/CNRS Editions, Paris.
127. Merzbacher [1970] : E. Merzbacher, *Quantum Mechanics*, John Wiley, New-York.
128. Messiah [1959] : A. Messiah, *Mécanique quantique*, Dunod, Paris.
129. Messiah et Greenberg [1964] : A. Messiah and O. Greenberg, Symmetrization postulate and its experimental verification, *Phys. Rev.* **136**, B248.
130. Mewes *et al.* [1996] : M. Mewes *et al.*, Bose-Einstein condensation in a tightly confining magnetic trap, *Phys. Rev. Lett.* **77**, 416.
131. Mohiden et Roy [1998] : U. Mohiden and A. Roy, Precision measurements of the Casimir force from 0.1 to 0.9 μm, *Phys. Rev. Lett.* **81**, 4549.
132. Nielsen and Chuang [2001] : M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge.
133. Norsen [2006] : T. Norsen, Bell locality and the non-local character of nature, Found. Phys. Lett. **19**, 633.
134. Nozières et Pines [1990] : Ph. Nozières and D. Pines, *The theory of quantum liquids*. Vol II, Addidon-Wesley, New-York.
135. Olmschenk *et al.* [2009] : S. Olmschenk, D. Matsukevitch, P. Maunz, D. Hayes, L. Duan and C. Monroe, Quantum teleportation between distant matter qubits, *Science* **323**, 486.
136. Omnès [1994] : R. Omnès, *The interpretation of quantum mechanics*, Princeton University Press, Princeton.
137. Parisi [1989] : G. Parisi, *Statistical field theory*, Addison-Wesly, New-York.

138. Peres [1980] : A. Peres, Non exponential decay law, *Ann. Phys. (N.Y.)*, **129**, 33.

139. Peres [1993] : A. Peres, *Quantum Theory, Concepts and Methods*, Kluwer, Boston.

140. Perkins [2000] : D. Perkins, *An Introduction to High Energy Physics*, 4è édition, Cambridge University Press, Cambridge.

141. Pethick et Smith [2001] : C. Pethick and H. Smith *Bose-Einstein condensates of dilute gases*, Cambridge University Press, Cambridge

142. Pines et Nozières [1989] : D. Pines and Ph. Nozières, *The theory of quantum liquids.* Vol I, Addidon-Wesley, New-York.

143. Pocholle [2005] : J.-P. Pocholle, De l'émission stimulée d'Einstein aux lasers d'aujourd'hui, dans *Einstein aujourd'hui*, p. 129, M. Le Bellac et M. Leduc eds., EDP Sciences/Éditions du CNRS, Paris.

144. Popescu et Rohrlich [1987] : S. Popescu and F. Rohrlich, Quantum non locality as an axiom, *Found. Phys* **24**, 379-85.

145. Preskill [1999] : J. Preskill, Quantum computation, http://www.theory.caltech.edu/˜ preskill/

146. Raimond *et al.* [1997] : J.-M. Raimond, M. Brune and S. Haroche, Reversible decoherence of a mesoscopic superposition of field states, *Phys. Rev. Lett.* **79**, 1964.

147. Rauch et Werner [2000] : H. Rauch and S. Werner, *Neutron Interferometry*, Clarendon Press, Oxford.

148. Rich [2002] : J. Rich, *Principes de la cosmologie*, Éditions de l'École Polytechnique, Palaiseau.

149. Rieffel et Polak [2011] : E. Rieffel and W. Polak, *Quantum computing : a gentle introduction*, MIT Press, Cambridge, Ma.

150. Riesz et Nagy [1955] : F. Riesz et B. Nagy, *Leçons d'analyse fonctionnelle*, Gauthier-Villars, Paris.

151. Rodberg et Thaler [1967] : L. Rodberg and R. Thaler, *The quantum theory of scattering*, Academic Press, New-York.

152. Santori *et al.* [2002] : C. Santori, J. Vukovic, G. Solomon and Y. Yamamoto, Indistiguishable photons from a single photon device, *Nature* **419**, 594.

153. Scarani [2003] : V. Scarani, *Initiation à la physique quantique*, Vuibert, Paris.

154. Scarani [2011] : V. Scarani, *Quantum information : primitive notions and quantum correlations*, dans *Ultracold gases and quantum information*, p. 105, Oxford University Press, Oxford.

155. Scarani *et al.* [2005] : V. Scarani, S. Iblisdir, N. Gisin and A. Acin, Quantum cloning, *Rev. Mod. Phys.* **77**, 1225.

156. Scheidl *et al.* [2010] : T. Scheidl *et al.*, Violation of local realism with freedom of choice, *Proc. Nat. Sci. USA* **107**, 46.

157. Schlossauer [2004] : M. Schlossauer, Decoherence, the measurement problem and interpretations of quantum mechanics, *Rev. Mod. Phys.* **76**, 1267 (2004).

158. Schmid-Kaler *et al.* [2003] : F. Schmid-Kaler *et al.*, Realization of the Cirac-Zoller controlled-NOT gate, *Nature* **422**, 408.
159. Schrödinger [1952] : E. Schrödinger, *Br. J. Philos. Sc.*, **3**, 109.
160. Schulman [1981] : L. Schulman, *Techniques and Applications of Path Integration*, Wiley.
161. Schwarzschild [2012] : B. Schwarzschild, Time-reversal asymmetry in particle physics has finally been clearly seen, Physics Today **65**(11), 16.
162. Shapere et Wilczek [1989] : A. Shapere and F. Wilczek eds., *Geometric phases in physics*, World Scientific.
163. Scully *et al.* [1989] : M. Scully, B. Englert and J. Schwinger, Spin coherence and Humpty-Dumpty III : the effect of observation, *Phys. Rev.* **A 40**, 1775.
164. Scully et Zubairy [1997] : M. Scully and M. Zubairy, *Quantum Optics*, Cambridge University Press, Cambridge.
165. Shimizu *et al.* [1992] : F. Shimizu, K. Shimizu et H. Takuma, Double-slit interference with ultracold metastable neon atoms, *Phys. Rev.* **A46**, R17 (1992).
166. Stöckman [1999] : H. Stöckman, *Quantum chaos*, Cambridge University Press, Cambridge.
167. Stolze et Suter [2004] : J. Stolze and D. Suter, *Quantum computing*, Wiley.
168. Streater et Wightman [1964] : R. Streater and A. Wightman, *PCT, spin and statistics and all that*, Benjamin, New-York.
169. Suter [1997] : D. Suter, *The Physics of Laser-Atom Interactions*, Cambridge University Press, Cambridge.
170. Tanzilli *et al.* [2002] : S. Tanzilli, W. Tittel, H. de Riedmatten, H. Zbinden, P. Baldi, M. de Micheli, D. Ostrowsky and N. Gisin, PPLN wave-guide for quantum communication, *Eur. Phys. Journ.* **D 18**, 155.
171. Taylor et Heinonen [2002] : Ph. Taylor and O. Heinonen, *Condensed Matter Physics* Cambridge University Press, Cambridge.
172. Taylor et Wheeler [1963] : E. Taylor and J. Wheeler, *Space-Time Physics*, W. H. Freeman, New-York.
173. Tinkham [1964] : M. Tinkham, *Group Theory and Quantum Mechanics*, Mc Graw Hill, New-York.
174. Tully [2011] : C. Tully, *Elementary particles in a nutshell*, Princeton University Press, Princeton.
175. Vandersypen *et al.* [2001] : L. Vandersypen, M. Steffen, G. Breyta, C. Yannoni, M. Sherwood and I. Chuang, Experimental realization of quantum Shor's factoring algorithm using nuclear magnetic resonance, *Nature* **414**, 883.
176. Wark [2005] : D. Wark, Neutrinos : ghosts of matter, *Physics World* **18**(6), June.
177. Weihs *et al.* [1998] : G. Weihs *et al.*, Violation of Bell's inequalities under strict locality conditions, *Phys. Rev. Lett.* **81**, 5039.
178. Weinberg [1978] : S. Weinberg *Les trois premières minutes de l'Univers*, Éditions du Seuil, Paris (1978).
179. Weinberg [1995] : S. Weinberg, *The Quantum Theory of Fields*, Cambridge University Press, Cambridge.

180. Werner *et al.* [1975] : S. Werner, R. Colella, A. Overhauser and C. Eagen, Observation of the phase-shift of a neutron due to precession in a magnetic field, *Phys. Rev. Lett.* **35**, 1053.

181. Wichman[1974] : E. Wichman, *Physique quantique*, cours de physique de Berkeley volume 5, Armand Colin, Paris.

182. Wightman [1960] : A. Wightman *L'invariance dans la mécanique quantique relativiste* dans *Relations de dispersion et particules élémentaires* p. 159, Hermann, Paris.

183. Wigner [1970] : E. Wigner, On hidden variables and quantum mechanical probabilities, *Am. J. Phys.* **38**, 1005.

184. Wheeler et Zurek [1983] : J-A Wheeler and W. Zurek eds., *Quantum Theory and Measurement*, Princeton University Press, Princeton.

185. Wootters et Zurek [1979] : W. Wootters and W. Zurek eds., Complementarity in the double-slit experiment : quantum non separability and a quantitative statement of Bohr's principle, *Phys.Rev.* **D19**, 473.

186. Wu et Shaknow [1950] : C. Wu and I. Shaknow, The angular correlations of scattered annihilation radiation, *Phys. Rev.* **77**, 136.

187. Zee [2010] : A. Zee, *Quantum field theory in a nutshell*, Princeton University Press, Princeton.

188. Zeh [1970] : H. Zeh, On the interpretation of measurement in quantum theory, *Found. Phys.* **1**, 69.

189. Zeilinger [1981] : A. Zeilinger, General properties of lossless beam splitters in interferometry, *Am. J. Phys.* **49**, 882.

190. Zeilinger [1999] : A. Zeilinger, Experiments and the foundations of quantum physics, *Rev. Mod. Phys.* **71**, S288.

191. Zeilinger [2000] : A. Zeilinger, La téléportation quantique, *Pour la Science* **272**, 36.

192. Zeilinger [2010] : A. Zeilinger, *Dance of the photons*, Farrar, Strauss and Giroux, New-York.

193. Zeilinger *et al.* [1988] : A. Zeilinger, R. Gähler, C. Shull, W. Treimer and W. Mampe, Single and double-slit diffraction of neutrons, *Rev. Mod. Phys.* **60**, 1067.

194. Zinn-Justin [1989] : J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, Oxford University Press, Oxford.

195. Zinn-Justin [2003] : J. Zinn-Justin, *Intégrale de chemin en mécanique quantique : introduction*, EDPSciences/CNRS Éditions.

196. Zurek [1991] : W. Zurek, Decoherence and the transition from quantum to classical, *Physics Today*, October 1991, p. 36.

197. Zurek [2003] : W. H. Zurek, Decoherence, einselection and the quantum origin of the classical, *Rev. Mod. Phys.* **75**, 715.

Index

D

E

F

P

Q

R

S

T

V

Z